School Closure in New York City

Megan Silander

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY
2012

ABSTRACT

School Closure in New York City

Megan Silander

School districts and states have increasingly abandoned traditional school reform efforts in favor of simply closing low-performing schools. This movement reflects growing frustration among policymakers with the disappointing effects of previous school improvement policies, and the view that some schools may simply lack the capacity to undertake meaningful improvements. This paper focuses on arguably the most aggressive school closure policies in the nation—those in New York City. Over the past decade, New York City has closed over 100 schools. Using a longitudinal database of students and schools, I explore the implementation and effects of closure and reconstitution of middle schools in New York City, and assess the links between school closure and student academic development and behavior.

My descriptive findings indicate that schools selected for closure have significantly lower school-average state test score exams and lower attendance rates compared to other middle schools for several years prior to closure, and that students who attend these schools are almost exclusively Hispanic and Black, more likely to come from low-income families, and more mobile than other middle school students in the district. I also find that students enter these middle schools already at a significant academic disadvantage. I examine characteristics of the reconstituted schools that replace the closed schools, and find that in terms of demographics, reconstituted schools enroll students similar to those served by the closed schools that they replaced. However, the reconstituted schools serve higher performing students with fewer absences and tardies in the year prior to enrolling in middle school.

To assess the impact of school closure on student academic outcomes, I use propensity-score matching within a difference-in-differences framework. I find a small, positive effect of school closure on student test scores and rates of absences. As a robustness check, I conduct a second set of analyses using student fixed-effects models that produced similar results: students learn slightly less at chronically underperforming schools, compared to what would have happened had they attended an alternate school.

School closure appears to be a somewhat effective in improving student academic outcomes. It is not clear, however, whether the policy is efficient given the small effects and the considerable disruption associated with the policy. Future research should examine the fiscal costs associated with closure, compared to costs of other policies with similar effects.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Background

Introduction

Over the past ten years, the New York City Department of Education has closed and reconstituted over 100 schools, and the pace of school closures is accelerating. In Spring 2012, the Department announced that it would close 42 schools the following school year alone. In support of the school closures, district administrators have cited chronically low test scores and graduation rates, as well as the necessity to demonstrate the very real consequences of the district's accountability system (Otterman, 2011). Mayor Bloomberg—a strong supporter of non-traditional approaches to school reform—has commented that those protesting the closures were either captives of nostalgia or were more concerned about jobs for teachers than the education of students attending poor-quality schools (Hernandez, 2011). Parents, students, teachers, alumni and local officials responded with heated anger and loud protests at the District's public school closure hearings (Otterman, 2011), and requested more academic and fiscal support to resurrect the schools that had served as the center of their communities, in some cases for decades (Robinson, 2010). Others argued that the fact that many of the schools slated for closure had been created under the existing Bloomberg administration suggested that the new closures and reconstitutions are unlikely to be any more successful (Rondinone, 2010). In spite of the controversy, District administrators have not wavered from the closure plans.

Opponents to the District's actions have charged that school closures are too costly and come at the expense of additional resources for struggling students (United Federation of Teachers, 2009); that school closure selection criteria are subjective and reflect political favoritism (Hernandez, 2009; Otterman, 2010c); that the reform is racist in its targeting of schools that enroll mainly Black and Hispanic students (Cramer, 2008a; Somerville, 2011); that

it does not take into account the additional needs and supports of the disproportionately disadvantaged students these schools currently serve (Somerville, 2011); that the reform is top-down and does not reflect the desires of local communities and families; that the measure is punitive rather than rehabilitative (Otterman, 2010a); and that it accelerates already substantial neighborhood gentrification, thus marginalizing such communities even further (Cramer, 2008b).

The District, however, counters that school closures represent the only means to dislodge elements of a moribund system entrenched within a failing bureaucracy (Hernandez, 2009). At the center of arguments supporting school closures is the contention that school closure results in improved student achievement. The dissertation examines the accuracy of this assertion within the context of New York City.

National Policy Context

Dozens of districts and states nationwide have implemented accountability policies that focus on school closures, including Chicago, which has closed 22 schools over the past ten years for poor academic performance (de la Torre & Gwynne, 2009) and Charlotte-Mecklenburg, which recently decided to close ten of its 178 schools (Helms & Price, 2010). Rhode Island (Kaye, 2010) and Kentucky (Maxwell, 2010) have also received national attention for their school closure efforts. But New York City is arguably the epicenter for this fundamentally different approach to education reform.

The shift from traditional school reforms toward more dramatic approaches is partly a response to policies initiated by the U.S. Department of Education. The federal Race to the Top (RttT) and School Improvement Grants (SIGs) competitions were designed to encourage states to take more forceful approaches to improving chronically low-performing schools (Otterman, 2010a). Three of the four federal policy options focus on school closure and reconstitution, and include closing a school entirely ("school closure"); replacing the principal and at least half the staff ("turnaround model"); closing and reopening a school as a charter school or under an educational management organization (the "restart model"); and replacing the principal while providing strategic school reform supports (the "transformational model;" US Department of Education [USDOE], 2010a; 2009). A recent analysis of the implementation of SIGs for the first and second cohorts of schools—in 2010 and 2011—suggests that although most schools are carrying out the school transformation model, another one-fifth of the 1,355 SIG-awarded schools are implementing the turnaround reform model, which most closely replicates closure and reconstitution (USGAO, 2012). In contrast, results from a survey of a nationally representative sample of school districts one year prior to the SIG program indicated that just

over ten percent had implemented a version of the federal turnaround reform model (Center on Education Policy, 2010), suggesting that the implementation of school closure and reconstitution has grown substantially over the past few years.

The No Child Left Behind Act (NCLB) also includes a series of incentives and sanctions for low-performing schools; the ultimate set of sanctions focuses on school "restructuring," under which school closure and reconstitution is an option (USDOE, 2008). Furthermore, the Obama Administration's Blueprint for re-authorization of the Elementary and Secondary Education Act proposes to continue this approach to school turnaround. Under the Blueprint, states must implement one of the four reform models for the bottom five percent of low-performing schools in each state (USDOE, 2010b).

In addition to these Federal mandates and guidelines, the budget deficits facing many districts and states have intensified school closure efforts. School districts, particularly those located in cities experiencing economic decline, increasingly are coupling school closure with fiscal considerations (Rich, 2012), including those in Boston (Kandile, 2011), Cleveland (Rich, 2012); Columbus (Rich, 2012), Detroit (Dawsey, 2012), Kansas City (Robertson, 2010), Oakland (Tucker, 2012), Providence (Beale, 2011), San Diego (Carless, 2012), South Bend (Kilbride, 2012) and Tucson (Wallace, 2011). Secretary Duncan has publicly encouraged state governors to consider closing low-performing schools as one method to simultaneously close achievement gaps and budget gaps (USDOE, 2011). Importantly, in these instances, school closures are justified in part by declining enrollment and new schools are not opened to replace the closed schools. In contrast, New York City is not obligated to close schools because of declining enrollments, which have remained relatively stable over the past decade. Raising

student achievement is the only justification for closing the city's schools, making clarification of the impact of the policy on student achievement an even greater imperative.

School Closure and Reconstitution: Theory of Action

Closure as Threat

The growing national reliance on school closure, along with the controversies that often follow its implementation, call for a closer examination of the theories on which the policy rests. Supporters of school closure and reconstitution claim two theoretical benefits. First, the threat of reconstitution can serve as a source of motivation, encouraging school staff to work harder with the hope of avoiding closure. Indeed, increasing the pressure on all schools to improve student achievement is a central aim of the policy. Thus, implementation of school closure as a policy remedy of last resort might be essential to maintain the integrity of local accountability systems. Regardless of whether school closure is effective for school buildings that are closed, school closure arguably increases the effectiveness of the accountability system by demonstrating that the sanctions within the system are indeed a threat, thereby motivating other similarly struggling schools. This sentiment is reflected in comments the Secretary of Education made in September, 2010 on Face the Nation: "You have to couple that autonomy with real accountability. And you have to hold them accountable for results. And we have five-year performance contracts. I ran the Chicago public schools. I was lucky to start many successful charter schools, but I closed three for academic failure. And so you have to couple that autonomy with real accountability."

However, research on the relationship between accountability threats and student achievement is inconclusive. Findings on the impact of high-stakes testing and accountability systems suggest that these types of incentive and sanction schemes provide no consistent positive effects on student learning. Several studies of the impact of accountability policies have found

no impact on student achievement in Charlotte-Mecklenberg (Smith & Mickelson, 2000) and Dallas (Ladd, 1999) mixed results in Chicago (Jacob, 2005; Neal & Schanzenbach, 2008). Other studies that examined multiple states found positive impacts on student achievement on high-stakes exams, particularly for those students already close to passing (Ballou & Springer, 2009) and positive impacts on low-stakes exams for all students, not just those close to passing (Reback, Rockoff & Schwartz, 2011). As these differential impacts suggest, some evidence indicates that high-stakes accountability systems caused schools to game the system. Specifically, schools focus their efforts on those students who are on the margin of passing and thus would make the greatest contribution to the accountability rating, so that the lowest-achieving students evidence lower or no learning gains on average (Burgess et al., 2005; Neal & Schanzenbach, 2008; Reback, 2008).

Little research examines the threat of sanctions specifically; evidence that disentangles accountability threats from stigma and incentives suggests no clear impact of accountability system threats on student achievement. Findings from a study of the NYC accountability system, in which schools that receive the lowest accountability ratings face possible removal of the school administration or closure and reconstitution, suggests that these threats result in improved student achievement. Students in schools that received the lowest accountability grades earned higher mathematics and English scores in the following year, compared to similar students in schools with very slightly higher accountability ratings, suggesting that accountability threats caused an increase in student learning (Rockoff & Turner, 2008). However, the study did not disentangle the impact of stigma from threat of closure or leadership removal. A related study in Florida found no impacts of accountability threats similar to closure, such as the threat of voucher assignment, on student achievement in English, and small positive impacts on

mathematics achievement, but only for students in the high-stakes grades being tested. Moreover, the results indicated that these gains were due to the stigma of the low accountability grade rather than the threat of vouchers (Figlio & Rouse, 2005). Similarly, examination of North Carolina's school-level NCLB accountability sanctions found no effect of the threat of a sanction, nor of the sanction itself (Ahn & Vigdor, 2009).

Benefits of Reconstitution

In addition to the use of closure to maintain the integrity of the accountability system, benefits may also flow from reconstitution itself, as new, supposedly more talented staff implement more effective pedagogical methods (Hess, 2003; Rice & Malen, 2003). School closure policies operate on the basic assumption that some schools simply lack the capacity to improve, a notion supported in part by the fact that schools are typically closed only after several years of mandated and ultimately unsuccessful reform efforts. This perspective represents a dramatic shift away from traditional conceptions of school reform, in that individual schools are sacrificed in service to what is perceived as the greater good. In this sense, the focus of reform is the local educational system, not individual schools.

Reforming Individuals

Human capital. Implicit also in the use of school closure as a reform is the assumption that the primary problem with failing schools is that the adults who work in them lack fundamental skills and dispositions that are fixed rather than fluid (Cohen, 2007; Rice & Malen, 2003). Hence, professional development is not viewed as an effective approach to school reform. Rather, reconstitution is predicated on a simple theory of action: only a more coherent and competent staff will improve student outcomes. How one develops such staff is not a concern. What matters is gathering adults who are able to address the school's needs, who bring

new abilities, beliefs, vision and focus (Goldstein, Keleman & Koski, 1998; Rice & Malen, 2003; Ziebarth, 2002).

Some proponents of school reconstitution also view the policy as a means of implementing a market-based approach to school reform. These advocates suggest that the system of tenure and seniority inhibits the free market flow of teachers across schools. The teacher labor market has difficulty attracting high-quality individuals to the profession and even more difficulty removing ineffective teachers. Thus, the closure and reopening of a school serves to force the teachers and leadership in the closing school to re-join the labor market as they are "excessed." In this manner, school closures arguably reinstitute essential market mechanisms eliminated under teacher tenure rules. Empirical assessments of closure's theory of action, however, are not particularly conclusive.

Research clearly supports the links between high-quality teaching and increased student learning. Teacher quality is one of the strongest influences on learning, with stronger effects than other factors such as student background, class size, or class composition (Aaronson, Barrow & Sander, 2007; Rivkin, Hanushek & Kain, 2005; Rockoff, 2004; Wright, Horn & Sanders, 1997). Research on leadership and its impacts on school is significantly less conclusive than the research on teacher quality, mainly due to the methodological difficulties in disentangling leadership effects from other school effects (Branch, Hanushek & Rivkin, 2009). Nonetheless, the evidence at least suggests that school principal quality directly affects other school conditions that relate to student learning, including student absences and discipline and teacher turnover (Clark, Martorell, and Rockoff 2009; Ladd, 2009; Leithwood et al., 2004), and through these conditions and mechanisms, indirectly impacts student learning (Branch, Hanushek & Rivkin, 2009; Hallenger& Heck, 1996; Horng, Klasik & Loeb, 2009).

While ceding this point, critics of reconstitution note that these theoretical benefits of school closure are rarely manifest in practice. Namely, reconstitution does not in fact result in more competent staff. Virtually no studies allow for causal claims regarding the impact of reconstitution on teacher or administrator quality. The research that does exist identifies a number of problems in the implementation of closure policies related to teacher quality. For example, due to a lack of qualified applicants, combined with deficient hiring processes, reconstitution policies often fail to raise the quality of teaching staff (Hess, 2003; Rice and Malen, 2003; Scott, 2008). Despite the use of school closures as a means to circumvent rules regarding the removal of ineffective teachers, district turnaround efforts often continue to be stymied by collective bargaining agreements that do not allow for the reassignment of the most effective staff to these schools (Scott, 2008). One study reports that almost three-quarters of newly hired staff in several reconstituted schools had less than one year of teaching experience (Rice & Malen, 2003), a characteristic associated with lower teacher quality (Aaronson, Barrow & Sander, 2007; Rivkin, Hanushek & Kain, 2005). Another study examining the implementation of school reform under the federal SIG program in California found that undertaking SIG-funded reforms, which included turnarounds, resulted in a decrease in average years of experience among school staff by over two years (Dee, 2012). The effect of these challenges, however, may attenuate over time (Hess, 2003).

Social capital. While human capital theory might call for the replacement of an entire school staff, social capital theory might call for similar but much less drastic measures. Under social capital theory, social relationships and structures facilitate—or constrain—productive activity, expertise, access to resources and the quality of these resources (Coleman, 1988; Dika & Singh, 2002). Although social capital theory applied to education frequently relates to the role

of the family and student social capital (Dika & Singh, 2002), it also can apply to school communities, as the relationships among teachers, administrators and other school staff facilitate instructional quality and access to institutional resources through shared norms about obligations and expectations, such as mutual support and reciprocity, levels of trust and access to expertise and information (Coleman, 1988; Dika & Singh, 2002; Penuel, Riel, Krause & Frank, 2009). Schools with high levels of social capital, then, may be more likely to have shared expectations and goals and greater capacity to work towards these shared goals through high levels of collaboration and collective problem-solving and the sharing of expertise and resources to improve instructional practice (Israel & Beaulieu, 2001; Penuel et al., 2009).

Social capital theory is similar to accountability incentive theories in that it attributes lack of school improvement to a lack of collective action (Dee, 2012), but has slightly different implications for how to increase collective action by reforming structures within the school site rather than outside the school community. Social capital theory suggests that moving teachers in a chronically underperforming to other schools with more productive and collaborative climates would improve their instruction and performance, while under human capital theory, we would expect these teachers to perform just as poorly in a new setting. While human capital theory calls for the replacement of the entire school staff with more competent individuals, social capital theory might require the replacement of a school's leadership or a smaller threshold of staff to allow for reshaping of the structure and processes to support social capital development. The minimum threshold could be close to the 50 percent staff replacement requirement of the federal school turnaround approach, or could be a smaller threshold limited to replacement of the principal, an approach that would fit under the federal transformation model.

However, there is little evidence regarding a causal connection between increased student achievement and reforms that target increased teacher collaboration through such structures as teacher teaming or professional learning communities, although the cross-sectional studies provide some suggestive evidence that efforts to increase social capital within a school could be effective (Corcoran & Silander, 2008; Lee & Smith, 1996; Supovitz, 2002). For example, one study found a positive relationship between student learning and levels of both collective responsibility for student learning and cooperation among high school staff. However, as this study is cross-sectional it does not support the conclusion that implementing a reform to increase collaboration and collective responsibility for student learning necessarily would impact student learning—that is, teacher beliefs about the limitations of students' ability to learn might not be alterable (Lee & Smith, 1996). Thus implications for closure and reconstitution versus the less dramatic school transformation options are unclear, particularly given that these studies tend to focus on the transformation of school social capital with existing school staff and few if any studies examine replacing school staff to develop more productive social capital.

Professional development. Opponents of reconstitution further contend that the reform is an overly simplistic approach that does not take into account the skills required to improve instruction (Ziebarth, 2002), and further, that school closures and the efforts required to hire new staff may detract from instructional planning (Scott, 2008). During the first few years of operation, tasks associated with planning and developing new policies, procedures, establishing school and professional development programs, ordering supplies and staff recruiting and hiring serve as a significant challenge for new school staff (Young et al., 2009). Further, new schools tend not to evidence large growth in student achievement in their early years of operation (Bifulco & Ladd, 2004; Gronberg & Jansen, 2001; Young et al., 2009).

Another set of critics of the current NCLB provisions and accountability structure suggest that failing schools need greater capacity to implement good instruction, and that rather than replacing teachers, districts need to provide instructional support for the existing teachers in the underperforming schools. In contrast to approaches to school reform in which teacher quality is understood as fixed, this alternative perspective argues for strengthening schools by enhancing teacher knowledge and skill (see Darling-Hammond, 2000; Ingersoll, 1999). Elmore (2006) characterizes the differences in these two competing approaches as "extracting" knowledge versus "developing" knowledge. Thus one could characterize reconstitution as an extracting approach; implicit is the assumption that teachers are either qualified or unqualified, and effective schools are able to extract that knowledge. In contrast, Elmore (2006) argues that under the development approach to school improvement, one assumes a base knowledge, but school organizations must work to continue to develop that knowledge to continue to provide effective instruction. In this manner, high-quality teaching is conceived as a technology, rather than a fixed attribute.

A central question, however, is the extent to which the instructional capacities of school-based staff can be improved. Research on the effectiveness of targeted professional development efforts is mixed, with few studies indicating that small-scale professional development efforts can increase student achievement (Carpenter, Fennema, Peterson, Chiang, and Loef, 1989; Desimone, 2009; Wayne, Yoon, Zhu, Cronen, & Garet, 2008), while more large-scale professional development programs with multiple trainers tend to have no evidence of impact on student learning (Garet et al., 2008; Glazerman et al., 2008). This variation in impact, combined with the fact that none of the studies that evidence positive impacts on student learning focus specifically on the impact of professional development efforts in our lowest-performing schools,

calls into question the use of professional development as a sole remedy for chronically underperforming schools. Moreover, the difficulties and inconsistencies in the effects of these types of programs raise the question of whether whole-school improvement is possible. I turn now to the broad body of literature that addresses this question.

Institutional Change

Advocates of school closure have suggested that the reform may serve as an external shock that dislodges ineffective school instructional practices and cultures in a way that other mandates are unable to do. One district administrator, in discussing New York City's approach to school closure and reconstitution describes the reform as targeting the school as an organization: "the question was not only changing the adults and the management of the school – the question was taking the opportunity to create school structures and enrollment patterns that were more likely to lead to success…. It was the targeting of dysfunctional school organizations…where we could wipe the proverbial slate clean so that we could begin the process of reconstitution" (Harries, 2009).

From an institutional theory standpoint, traditional school reform efforts fail because they do not take into account the institutional environment of schools—specifically, the notorious impenetrability of school cultures and their ability to resist change. Instruction, and the policies and structures that govern it, are "loosely coupled" (Meyer & Rowan, 1977, 1978), such that schools satisfy bureaucratic demands by making symbolic changes rather than investing in efforts to transform classroom practice. The challenge is that bureaucratic requirements are often unrelated to organizational effectiveness (Honig & Hatch, 2004; Meyer & Rowan, 1977, 1978; Weick, 1976). Interpreted in this light, school closures represent an attempt to circumvent

existing school cultures by creating new organizations that are more tightly linked to the delivery of instruction.

The generally weak track record of school reforms lends some support to this view. One meta-analysis of school reform initiatives reported that the success rate, defined as sufficient improvement in standardized test scores to meet state standards, stood at around fifty percent (Brady, 2003). This finding is reflected in more recent studies of improvements in average school achievement over time. For example, not taking into account changes in student characteristics, approximately two-thirds of California schools in the lowest quartile of average performance in 1989 remained in the lowest quartile twenty years later (Loveless, 2010). A related study examining both low-performing charter and traditional public school achievement from 2002 to 2008 in ten states found that approximately twenty percent of these schools were able to improve student achievement to rise above their state's bottom quartile of proficiency, and that charter schools were no more likely than traditional public schools to improve (Stuit, 2010).

The results of organizational turnaround efforts in other sectors provide a similarly dreary picture (Kowal & Hassel, 2005). For example, Kotter (1995) asserted that 70 percent of turnaround efforts in the private sector fail. Studies of specific private sector turnaround approaches such as Total Quality Management and Business Process Reengineering echo these findings (Hess & Gift, 2009). The general failure of school reforms to improve chronically under-performing schools, coupled with similar disappointments in other sectors, suggests that school improvement may require abandoning traditional within-the-system approaches to school reform for more dramatic disruptions of school instructional practices and cultures.

Research on the development of small schools and small learning communities (SLCs) within schools provides some support for the creation of new schools rather than the restructuring of existing schools. Justifications for creating smaller schools and SLCs are commonly based on the premise that smaller schools can offer better guidance and personal attention for students, decrease disciplinary problems and create a safer school environment, increase teacher empowerment, leadership, and collaboration within and across disciplines, promote more efficient administration via a simpler administrative hierarchy, and lead to a more responsive and focused curricula (Oxley, 1989; Raywid, 1995). Small learning communities (SLCs) have arisen out of the small schools movement as a feasible method with which to implement the small school model using existing structures (Springer, Houck, Ceperley & Hange, 2007; Lee & Ready, 2007; Raywid, 1995).

Research on the creation of small standalone schools suggests that small schools tend to have better academic outcomes, higher rates of attendance, more positive climates and fewer disciplinary problems compared to large schools (Barrow et al., 2010; Bloom et al., 2010; Fowler, 1992; Lee & Smith, 1995). In contrast, research evidence on the impact of SLCs suggests mixed effects (AIR & SRI, 2005; Herlihy & Kemple, 2004; Kemple, Herlihy & Smith 2004). These studies reported that restructuring schools into SLCs can result in structural changes unconnected to teaching and learning despite support from teachers and administration for the reform. For example, restructuring schools into SLCs does not necessarily result in increased teacher collaboration, and even if the restructuring fosters new practices, they do not necessarily lead to more innovative nor more effective teaching (Lee & Ready, 2007). The positive impacts of new small schools on learning, compared to the lack of impact, on average, of schools that are restructured into SLCs provides some suggestive, although not conclusive,

evidence that school closure and creation of new schools may be a more effective reform compared to restructuring or re-organizing an existing school.

Students, Families and Communities

Student composition. Critics of school reconstitution contend that one of the most disturbing aspects of the policy is that the new schools do not serve the same types of students that the closed school had served. These critics charge that the new schools do not serve the students in the community, or that the reconstituted schools select only higher-performing students (ACNY & AALDEF, 2009; Plasencia & Agarwal, 2009). For example, reform efforts in Chicago may have encouraged principals to push out students who were chronically absent (Hess, 2003). As such, already under-served communities may be left with even fewer educational resources. Furthermore, school closures may represent an added burden on students who are reassigned to schools outside of their neighborhood (Bellandi, 2009).

On the other hand, it is possible that such changes in enrollment patterns might positively alter the socio-demographic composition of schools. Although not always explicitly stated, New York City's reconstitution reforms are part of a larger effort to move away from neighborhood schools to what the administration perceives as more equitable district-wide choice and enrollment patterns (Harries, 2009; NYC DOE official, personal communication, October 2008). In short, reconstitution may serve as a means to lessen the high concentrations of poor and low-achieving students.

Research on the relationship between school composition and student learning provides some support for this potential impact of reconstitution. Specifically, higher levels of aggregate peer achievement are related to increased student learning, and this impact is stronger for students with lower levels of achievement (Barber, 1961; Borman & Dowling, 2010; Finn, 1987;

Hanushek, Kain, Markman and Rivkin, 2003; Summer & Wolfe, 1977; Zimmer & Toma, 2000). Research on school choice suggests that when students move to a school with higher-average performance levels, students tend to learn more, on average (Hastings & Weinstein, 2008). However, it is important to note that little research exists on impact of school reconstitution on the composition of a school's student body, specifically. Moreover, anecdotal evidence suggests that closure may result in greater concentration of low-achieving students in schools nearby the closed school as these schools absorb students who would otherwise have attended the failing school (Freedman, 2008; Robinson, 2011). Addressing these unresolved questions is a central focus of my dissertation.

Parent and community involvement. Although the above research on student body composition suggests that altering enrollment patterns might result in greater equity, others contend that closures may sever important links between schools and the surrounding neighborhoods. Opponents of school closure argue that local connections are important in ensuring parent involvement in schools and in creating and tailoring a school environment and program that is more responsive to local needs and values, and thus more effective in serving students. On an individual level, connections between parents and their children's schools, such as communication with staff and attendance at school events, positively impacts student achievement (Epstein & Sanders, 2002; Grolnick & Slowiaczek, 1994; Hill & Tyson, 2009; Jeynes, 2005; 2007; Lareau, 1987). Conversely, research on student mobility is able to link at least part of the detrimental effect of mobility on achievement to declines in social relationships or social capital (Pribesh & Downy, 1999) as students shifting to new schools following closure experience weaker relationships with teachers (Kirshner, Gaertner & Pozzoboni, 2010).

In addition to individual family involvement, at the collective level, parent collaboration and participation in local school councils and in local decision-making might serve to increase the accountability and effectiveness of the school administration and school bureaucracies and make reforms more sustainable (Fine, 1993; Fung, 2004; Henderson & Mapp, 2002; Henig et al., 2011). Participation in decision-making processes might also serve to increase parent engagement and increase parental expectations. However, it is difficult to definitively connect school closures and reconstitution to decreases in individual or collective parental involvement, particularly in urban areas where the alternate receiving schools are often a few blocks from each other, and given that parent involvement is likely already quite low in the under-performing schools selected for closure. Moreover, these concerns may be less salient in New York City, given its somewhat unique approach to school closures.

Impact of Closure and Reconstitution on Student Learning

Despite the considerable controversy surrounding the reform, its dramatic effects on students, school staff, and communities, and its increasing popularity, surprisingly little research has assessed the causal impacts of school closures on student outcomes (Rice & Malen, 2003). Those studies that did so were mainly small descriptive case studies (Galleta & Ayala, 2008; Rice & Malen, 2003; Goldstein, Keleman & Koski, 1998) or observational studies (Scott, 2008; Hess, 2003), and often covered only a few years (Hess, 2003; Rice & Malen, 2003). More recent studies suggest mixed results. A study of the implementation of the federal SIG program in California, which includes both the federal turnaround and transformation model to school reform, found a small positive effect on student test scores for the lowest-achieving schools, with higher impacts for schools using the school turnaround model. The study did not find significant effects for "higher" low-performing schools (i.e., "persistently low-achieving schools" that did

not meet the lowest-achieving five-percent threshold but did meet the "lack of progress" threshold as defined under the SIG guidelines; Dee, 2012).

Additionally, a few studies have examined district-level school closure reforms in detail. One investigation of Chicago K-8 schools found that students leaving closed schools tended to transfer to equally poorly performing schools. Examination of student learning one year later suggested no impact on student learning (de la Torre & Gwynn, 2009). Student learning varied by the quality of the receiving school, with students transferring to schools with higher average achievement learning more than they would have had they remained in the closed school. A small-scale study of the closure of one school in another urban district found a decline in student test scores in the two years following closure, and a decline in the probability of graduation (Kirshner, Gaertner & Pozzoboni, 2010). Note, however, that the policy in both studies was to close schools and move students immediately rather than allowing for students to phase out with the school as New York's policy is designed to do. Thus, it is not clear how applicable these conclusions are to New York, where problems related to student mobility and efforts to establish ties to new adults may be less relevant. My dissertation is an attempt to address this major weakness in the literature.

New York City Context

Before examining the implementation and impact of closure and reconstitution in New York City, it is important to understand the details of the policy and contexts governing and shaping the reform initiative, and to illuminate the specific policy mechanisms and conditions that may drive any impacts on student outcomes. These details may also assist in identifying whether it is possible to generalize this study's findings to other contexts.

New York City Public Schools

The New York City Department of Education governance structure is a form of mayoral control. In 2002, the state legislature granted Mayor Bloomberg control over the city school system. The same year, Bloomberg appointed Joel Klein as Chancellor. In January 2011, Cathleen Black succeeded Klein as Chancellor of the city schools; she was replaced in March 2011 by Dennis Walcott. Although I describe school closures from 1998 through 2008, my analysis of the impact of closure focuses on closures during the Bloomberg/Klein administration, which launched a series of district-wide school reform initiatives beginning in 2002-3, titled the Children First Initiative.

The Children First Initiative encompassed a series of sweeping, District-wide reforms. These efforts began with the establishment of mayoral control and the dismantling of the 32 local community school boards. A central focus of Children First was on devolving power from the central office and local community districts to principals and schools. Principals were granted greater authority over hiring processes, school budgets, instructional programs, and professional development. Rather than report to local district superintendents, schools selected from twelve "School Support Organizations" that were responsible for meeting schools' professional development needs and other support. The District granted this increased autonomy to schools in

exchange for increased accountability for meeting specific performance goals. The Department also introduced a number of a number of structures and processes to support accountability, including student, teacher and parent surveys of the school environment, annual progress reports and more in-depth school quality reviews, the establishment of inquiry teams of teachers and staff at each school to examine student data and chart improvement strategies, and the development of a comprehensive electronic data system (O'Day, Bitter & Talbert, 2011). Also during this time, the administration began implementing its central school turnaround policy, the New School Initiative.

Federal and State School Turnaround Policies

Governing the DOE's approach to school turnaround are Federal and New York State policies and initiatives. Specifically, the No Child Left Behind (NCLB) act requires that states set annual performance standards and track school progress towards meeting the standards based on student performance on annual state assessments. Schools must make "adequate yearly progress" (AYP) in the percentage of students who pass these tests, and states must identify Title I schools that fail to make AYP. NCLB's ultimate sanction for schools that continue to fail to meet AYP for the fifth year is "restructuring," which focuses on changing the governance and staffing of the school. The provisions under this sanction call for districts to implement one of five restructuring options: close and re-open the school as a charter school; arrange for a state to take over the direct operation of the school; replace the entire school staff, including the principal; turn management of the school over to a private agency; restructure school governance in accord with state direction; and a sixth option, loosely defined as undertaking "some other restructuring of school governance" (Scott, 2008).

New York State monitors school performance using the "Registration Review" process, first established in 1989. Under this state accountability policy, the state identifies schools not making adequate progress on state assessments; schools that are "farthest from meeting these minimum standards" and "judged by the Commissioner to be most in need of improvement" are placed on the Schools Under Registration Review (SURR) list and threatened with the possibility of losing registration; the final sanction for these schools is closure (NYSED, 2010a). Between 1989 and 2010, the state closed 70 schools across the state for lack of adequate progress; this number includes some but not all of the schools closed by New York City (NYSED, 2010b). Although the identification of the 'worst' schools in New York City sounds quite similar under federal and State policies, there is little overlap between those identified by New York State and through the federal accountability system (Stiefel, Bel Hadj Amor & Schwartz, 2005), highlighting the complications that may exist in implementing these policies.

In the past few years, New York State has heightened its focus on school closure in response to federal policy changes regarding school turnarounds, including the Race to the Top and School Improvement grant initiatives. In 2010, New York State broadened its list of schools to be closed to include the bottom five percent of schools in the state, identified as "Persistently Lowest Achieving" (PLA); all schools in this category (including SURR schools) must implement one of the federal government's four turnaround strategies, one option of which is closure and reconstitution. The state assigns schools to the PLA category that are in the restructuring phase of the accountability system and that did not meet a minimum performance level the year prior and did not make a minimum point-gain over three years prior to the designation. Additionally, high schools with graduation rates below 60 percent for three years prior to the current school year were also added to the list (NYSED, 2010c).

New York City's Approach to School Turnarounds

New York City has taken a more aggressive approach to school turnarounds than called for by state guidelines, often preemptively restructuring schools. District administrators describe NYC's approach to school turnarounds during Chancellor Klein's tenure as centered on closure and restart (Harries, 2009). The city's central turnaround effort, the New School Initiative, resulted in the closure of over 80 schools and the creation of 335 new traditional schools and over 86 new charter schools between 2002 and 2009. The majority of the schools that have been closed and re-opened were high schools, although the district also closed 31 middle schools and opened 66 new middle schools during this time period. The district's approach to turnarounds has focused on changing the staffing and management of chronically low-performing schools, and ostensibly altering both school structures and enrollment patterns by replacing large failing schools with small schools and reducing the concentration of low-achieving students within a school (Harries, 2009).

The city bases decisions to close schools primarily on academic achievement (Harries, 2009; Independent Budget Office [IBO], 2010). The current closure process includes identifying schools that receive a grade of a D or lower on the most recent school progress report or three consecutive grades of C, or that have been rated as "below proficient" on the most recent School Quality Review. Because these criteria include a large number of schools, the DOE narrows the list of schools by eliminating from selection new schools, schools with a "well-developed" rating on their School Quality Review, elementary or middle schools that have test scores above their community district average, and high schools with graduation rates above the city average (IBO, 2010). However, the decision is not based solely on accountability data. The DOE's decision to close a school is also based on whether enrollment has declined at the school (a potential

indicator of parent or student satisfaction), data on school staff and leadership and community needs, and input from superintendents and school network leaders (Harries, 2009; IBO, 2010).

Specific school closure policies prior to the Bloomberg administration are more difficult to disentangle. The closure policies do not appear to have been as systemic in nature as those in place during the Bloomberg and Klein administration. Two different chancellors served between 1998 and 2002, and New York City school governance was less centralized prior to 2002. One result was that Local District Superintendents had more control over school closures. For example, Brooklyn middle school closures in 2000, at which time the District was under Chancellor Harold Levy's tenure (2000-2002), were led and carried out by Local Superintendent Carmen Fariña due to declining enrollment and a budget shortfall (Gittrich, 2001). At least three middle school closures in 2001 were at the behest of New York State; the public rationale for closure focused on the fact that these schools had languished on the New York State SURR list for over a decade (Shin, 2000). Closures under Chancellor Rudy Crew in 1998 and 1999 appeared to have been driven by the Chancellor and not the Local Superintendents (who were themselves targeted by the Chancellor for removal). Chancellor Crew cited low test scores, as well as numerous years spent on the New York State SURR list, as justification for these closures (Hartocollis & Holloway, 1999). Closures in 1997 also were directed by Chancellor Crew (although there is some evidence that Local District Superintendents exercised some power in preventing, if not implementing, school closure during this time) and conducted in response to the State's demands (Hartocollis, 1997; Sorensen & Goodman, 1997), while the 1998 and 1999 closures appear to have been pre-emptively conducted by the city.

Once the city announces closure decisions, parents, students, community members and some local politicians usually respond with opposition, which appears to have grown stronger

and more acrimonious over the past few years. For example, one community group sued the city to allow them to site their protests on the sidewalk in front of the mayor's residence (Otterman, 2010). During another protest, twenty-four protestors, including two city councilmembers, were arrested (Christ, 2011). This increased opposition may stem from several factors: the district has accelerated closures each year; the schools closed in later rounds are less clearly dysfunctional (Robinson, 2011); and a growing perception that district policy decisions are unilateral.

Despite this recent vocal opposition, only in a few cases have groups managed to alter closure decisions. In Winter 2010, the NAACP, teachers' union, local politicians and other community members filed a successful lawsuit to postpone school closures that year, based on a technicality related to a new state regulation requiring the district to provide an educational impact statement and extensive advance notice of closures. However, the city reintroduced the decision to close schools the following year (Walz, Arp & Phillips, 2011). In response the UFT and NAACP filed a second lawsuit against the closures, but was unsuccessful (Cramer, 2011). In 2012, the UFT brought a third lawsuit, which was successful in reversing the District's proposed turnaround—a process similar to the District's approach to closure in that it would require firing the principal and more than half the staff, but would not displace students—of another 24 schools in the 2012-13 school year (Baker, 2012).

Critics of the city's selection process suggest that the schools identified for closure serve students who have much greater needs than students at other schools. For example, a 2010 Independent Budget Office report found that students in high schools identified for closure were more likely to require special education and English-language learner services, come from low-income families, live in temporary housing, have lower eighth-grade test scores and to have been

retained, compared to high schools that were not selected for closure. However, these disparities were not found in the middle and elementary school comparisons (IBO, 2010).

Phase-out. New York does not close a "failing" school and immediately open another. Rather, schools are, with a few exceptions, phased out and phased in over several years. For example, once a school is slated to be closed, the school will cease admitting new students that next fall, but allow currently enrolled students to continue, closing once the last group of admitted students has graduated. DOE administrators suggest that the phase-out process allows for more personal attention to students as the schools get smaller and that students in phasing-out schools are more likely to graduate than students who attended the school prior to the phase-out process (Cramer, 2011; Fertig, 2011). Critics suggest, however, that as the phasing out schools lose facilities such as libraries, drop extra-curricular activities and excess teachers, that the city has underestimated the drop-out rates during phase-out (Rowland, 2011).

Reconstitution. As schools phase-out, new schools typically phase-in within the same building. Many of these new schools are smaller than the closed school, therefore multiple schools may open in a building at the same time. In opening the new schools, rather than prescribing instructional practices, the district has emphasized the need for a coherent instructional plan and evidence of capacity to implement it from the new school founders and administration. To support leadership capacity, the district developed the New School Development and New School Intensive Professional Development Academies, in partnership with the New York City Leadership Academy, to provide future principals of the new schools with professional development and support in planning the turnaround process. Additionally, the district has planned for the phasing in or growth of a new school over several years to allow time for new principals to develop capacity to lead. Most of the new small schools are supported by

partnerships with educational management or other nonprofit and community organizations that hold expertise in new school development. The district provided schools with additional start-up funds as well (Harries, 2009).

Enrollment processes in these new small schools are the same for any district school. Most elementary school students attend local zoned schools, although some gifted and talented and other specialized schools draw students from a whole community district. Enrollment processes in middle schools vary. Some middle schools are zoned, but many middle schools enroll any student within the community district or in some cases, borough-wide. The unzoned schools enroll students through a special application or lottery. At the high school level, all students must select a school to attend and enrollment is district-wide. Processes vary from the most selective schools, which require specific achievement levels, to limited unscreened schools, which give preference to students who attend a school information session, to unscreened schools in which students are randomly selected by a computer, sometimes based on ensuring a diverse student body. Finally, the district also sites charter schools in the buildings formerly occupied by the closed schools. In the charter schools, enrollment is based on lottery, and is not circumscribed by catchment area (Inside Schools, 2011).

Anecdotal reports suggest that nearby schools often receive an influx of high-need students when a neighboring school is closed, many of whom enter with particularly low test scores and/or behavior issues because the newly opened schools will not accept them (Freedman, 2008; Robinson, 2011). A report conducted by the Center for New York City Affairs found that many remaining comprehensive high schools experienced an influx of students between 2003 and 2008 (Hemphill, Nauer, Zelon & Jacobs, 2009). However, because the report provides no comparison to trends in similar schools not located near closures or from where the influx of

students are enrolling, it is impossible to know whether the increases in attendance are due to the closures specifically. Additionally, prior to 2007, for the first three years of operation the newly reconstituted schools were not required to enroll students that required special education or English language learner services, under the theory that they would need to spend the first few years of operation developing sufficient capacity to serve the additional needs of students in these categories. However, in response to a complaint filed with the U.S. Office for Civil Rights and criticisms that these newly reconstituted middle and high schools were not serving the same population of students that attended the closed school, the city changed its policy for the new schools in 2007, requiring that new schools accept ELL and special education students when they open, rather than delaying serving these types of students (Gewertz, 2006; Gootman, 2006). The lack of evidence surrounding these competing claims calls for more research on the implementation and impact of school closures as a reform for low-performing schools.

Research Focus and Guiding Questions

Despite the considerable controversy surrounding the reform, surprisingly little research has assessed the impact of school closure and reconstitution on students and schools. I begin by examining issues related to implementation of the policy in the first results chapter. Which schools are selected for closure and how can we characterize the types of students and communities who experience it? In that chapter, I explore the extent to which the characteristics of reconstituted schools actually differ from the schools they replaced. In the second results chapter, I describe the implementation of the school phase-out process. I examine impact in the final results chapter. Using an expansive ten-year longitudinal database of students and schools in New York City, I assess the links between school closure and student academic development.

For the impact analysis, there are a number of conceivable impacts, treatment groups, and

counterfactuals. I focus specifically on comparing outcomes for two groups of students: students who attend the to-be-closed schools *before* school closure is announced, compared to students who would have attended the to-be-closed schools had they not closed and who instead attend the schools that reconstituted the closed schools or other neighboring schools. That is, I compare the academic development of students who *would have* gone to the closed schools to that of students who *did* go to the closed, chronically failing schools, before these schools were identified for closure and ceased to enroll students. I do not explore, for example, the impact of phasing out a school for students who attended the to-be-closed schools at the time of the closure announcement and later.

Chapter 2: Data and Methods

Data

My dissertation employs longitudinal New York City Department of Education administrative data on New York City students from 1999 to 2009, supplemented with school-level data from the New York State Department of Education and the Common Core of Data from the National Center for Education Statistics. New York City is the nation's largest public school district, serving approximately 1.1 million students in just under 1,700 schools. The city serves a diverse group of students: 31% are black, 40% Hispanic, 14% white, and 15% other minorities, including Asian and multi-racial students (NYSED, 2011). Just under half of the student population (42%) speaks a language other than English at home (NYC DOE, 2007).

I focus specifically on middle schools students. My base sample includes students with at least one test score for a middle school grade (sixth, seventh or eighth grade; n=1,066,175). Because middle school students are more likely to have both mathematics and English test scores available annually several years prior to entering middle school and are tested every year during middle school, this age group lends itself well to a longitudinal analysis of the impact of school closure on student learning. High school students, in contrast, are not required to take English and mathematics assessments annually, and elementary school students are not assessed until third grade. Moreover, attendance zones at the middle school level are more circumscribed by the 32 geographic community school districts, compared to high school enrollment, which is city-wide.[1] Slightly more constrained attendance patterns allow for the establishment of more plausible counterfactuals of the schools students would have attended had their school not

[1]The New York City School District is divided into 32 community school districts across the five boroughs. These community districts constrain attendance zones for middle and elementary school students so that, with the exception of a few specialized schools, middle and elementary schools only serve students from one of the 32 districts. The districts hold relatively limited administrative duties, related to zoning, supervision of school administrators and parent support.

closed. Additionally, in contrast to the high school reconstitution initiative, the District's approach to closing middle schools does not always include replacement with new small schools. In some cases, middle schools were replaced with schools of the same size. Thus the impact of small school reform potentially can be disentangled a little more clearly from closure. Finally, the lack of research on middle schools and reconstitution provides a compelling reason to examine the policy from the perspective of middle schools.

Measures

Academic outcomes. Student scores on annual New York State English and mathematics standardized assessments and student absences, as measured by number of days absent per semester, serve as my primary outcomes.

Standardized assessment outcomes. Beginning in 1999, students in grades 4 and 8 were administered state standardized assessments, including multiple choice and open-ended questions aligned to state standards in English Language Arts (ELA) and mathematics. In an effort to assess every student annually, the city supplemented the state assessments with city assessments for students in grades 3, 5, 6 and 7. These city assessments were aligned with the state standards but included only multiple choice questions (NYC DOE, 2009). Both of these state and city assessments are criterion-referenced, and produce IRT scale-scores that are vertically equated, allowing for comparisons within and across grades. However, in Spring 2006, New York State expanded the state assessment program to include assessment of all students in grades 3 through 8, thus replacing city assessments in grades 3,5, 6 and 7. The new 2006 New York State assessments are not vertically equated and do not allow for comparisons of student academic growth across grades and years. Rather, the assessment used in 2006 and later years allows comparisons between the same grades from year to year (CTB/McGraw Hill, 2006).

Additionally, the state's alteration of the 2006 test scale, content, scores and range do not allow the scores from 2006 and later to be compared to scores from 1999-2005 (NYC DOE, 2008). Therefore, to account for these assessment changes, scores are standardized within grade, year and subject. Because they are aligned to the state standards, I assume that despite changes in scaling and scores, the New York City and New York State assessments measure similar aptitudes and thus once standardized are comparable across years within in each subject and grade. Similar approaches have been taken by Boyd et al. (2008) and Schwartz and Steifel (2006).

Social and academic background. It is likely that students who attend schools that are closed are quite different from students who attend other district schools. Therefore, I take into account a number of student social and academic characteristics that are related to both attendance at a chronically underperforming school and student outcomes. These characteristics include the number of absences and tardiness incidents each school year. For the matching difference-in-differences and fixed effects models (described below), I also coded the attendance measures into five-level categorical measures by quantile, to allow for non-linear relationships. I measure race/ethnicity using a series of dummy-coded measures of whether the parents reported the child was American Indian, Asian or Pacific Islander, black non-Hispanic, white non-Hispanic, multi-racial, or the parent refused to report, using Hispanic as the uncoded comparison group in my analytic models. The models also include an indicator of whether the student ever received special education services (yes=1, no=0), whether a language other than English was the primary home language (yes=1, no=0), free/reduced lunch eligibility (yes=1, no=0), and whether the student was old for their grade either due to delayed school entry or retention, defined as 15 months older than the New York City minimum age for sixth grade at the time of

entering sixth grade for the first time (yes=1, no=0). I also account for student mobility, coded as the number of elementary schools a student attended.

School characteristics. School characteristics include year of closure, a measure that I calculated by identifying any schools that ceased to enroll students each year. I also confirmed this identification using either news reports, NYC DOE press releases, or on-line reviews from Inside Schools, a local organization that provides comprehensive evaluations of city schools. Other school characteristics include continuous measures of school size (number of students enrolled), percent of students eligible for free/reduced lunch, percent of students whose home language was not English, percent of students who attend three or more elementary schools, school-average achievement as measured by state and city assessments, and school-average absence and tardiness rates. I measure racial/ethnic composition using separate continuous measures of the percentage of Hispanic, black, white and Asian students in the school.

Missing Data

As with most large administrative datasets, students in my sample are missing data on particular measures, I relied on imputed data for the impact analyses but used on the existing administrative data and federal Common Core of Data from the National Center for Education Statistics for the descriptive school-level analyses. Tables 2-1 and 2-2 display the percentage of cases missing data for the cohorts of students in the impacts analyses, those students who were enrolled in sixth grade in 2001 and later years. Rates of missing data are higher for earlier cohorts.

Table 2-1
Percent of missing data for student demographic characteristics

	Percent missing
Race/ethnicity	4.5
Sex	10.0
Age	2.8
Free/reduced lunch	1.6

Table 2-2
Percent of missing data for student student academic characteristics

	Percent missing					
	Third grade	**Fourth grade**	**Fifth grade**	**Sixth grade**	**Seventh grade**	**Eighth grade**
ELA test scores[a]	27.9	23.3	18.6	14.6	26.3	36.5
Mathematics test scores[a]	22.9	19.1	15.3	11.1	22.9	33.9
Absences						
Fall	51.4	49.7	36.7	21.4	21.8	32.3
Spring	61.5	38.8	24.9	20.2	32.2	43.0
Tardies						
Fall	61.5	49.7	36.7	21.4	21.8	32.3
Spring	51.4	38.8	24.9	20.2	32.2	43.0

Special education status was a particularly problematic measure. The dataset included a two indicators of whether the student has an IEP from different administrative datasets, and the two indicators did not overlap entirely. Furthermore, these two indicators did not overlap entirely with other indirect indicators of special education status, such as whether the student received an IEP diploma, was eligible to take the New York State Alternative Assessment (NYSAA), or attended a special education school. Because of these problems, I coded students as having an IEP if they were coded positively for any of these indicators. Similarly, whether a student's home language was not English variable did not include a negative indicator: it was coded "Y" if a student's home language was not English, and blank otherwise. Therefore, missingness was not determined for these two measures.

It is likely that these data are not missing completely at random. For example, low-income students are more likely to move in and out of schools and the district and thus more likely to be missing test scores. Students whose families do not speak English or have limited literacy skills are also less likely to complete administrative questionnaires used to obtain demographic data such as gender, race/ethnicity, free/reduced lunch status. Simply dropping cases that are missing data would result in a dataset that is more advantaged than the New York City student population, and thus may bias my results. Additionally, because many students have at least some missing data, dropping these students and any of their existing measures results in the loss of information that could otherwise inform my analysis (Allison, 2002; Rubin, 1987).

In response, I employ multiple imputation using the "Amelia II" package available for R to generate five imputed datasets (Honaker, King & Blackwell, 2011). The multiple imputation approach to estimating missing data offers a number of benefits. The method draws from the full dataset and thus does not ignore what might be valuable information, unlike list-wise deletion, thus affording less biased estimates (Allison, 2002). Additionally, multiple imputation retains the relationships between variables within the data, but allows for a more streamlined method to impute multiple variables that have missing data simultaneously, rather than variable by variable. A third advantage is that multiple imputation accounts for the two types of uncertainty surrounding the missing data: model uncertainty and sampling uncertainty. A major difficulty, however, with this approach is that it may be difficult to implement using a large dataset with multiple missing values, particularly given computer memory limitations. Because of these limitations, I divided the data into 20 randomly generated groups and imputed each group separately, using in each an indicator of cohort (year the student first entered sixth grade) to allow the missing data patterns to vary across cohorts of students. I conduct matching and

estimate the effect of closure using the difference-in-differences and fixed effects approaches separately on each of the five imputed datasets, and then combine the results based on Rubin's (1987) approach. In my imputation models I included all variables that were used in the later analyses.

Methods

Describing Students and Schools

My initial analyses focus on the implementation of school closure and reconstitution. For these descriptive analyses, I include all data on closed middle schools from 1999 through 2008. I define middle schools as schools whose lowest grade is between four and seven and highest grade between five and nine. I conduct chi-square tests, t-tests and one-way analyses of variance (ANOVA) analyses to determine statistically significant relationships.

Estimating the Effects of School Closure on Student Learning

Because of the timing of school closures in New York City, combined with the periods captured by my data, I limit my impact analyses to middle schools (grades 6-8) whose last year of operation was 2006 (5 middle schools closed), 2007 (3 middle schools closed) or 2008 (6 middle schools closed). Figure 2-1 displays the various cohorts of students affected by the closure and phase-out process for each of the three years of closure in my analyses. I compare outcomes for students who did enroll in the to-be-closed schools to those who would have had the schools not begun to close. The treatment cohorts—students who attend and graduate from the chronically failing to-be-closed schools before closure is announced—are colored black. The control cohorts—students who attend an alternate receiving school after closure is announced and the to-be-closed schools cease admitting new students—are indicated by the dark-gray cells. These control students attend either the newly reconstituted schools that replace the to-be-closed

schools, or other neighboring schools. Students who attended the schools during the transitional, phase-out process (not included in this study) are indicated by the light-gray cells. I examine specifically the impact of attending a school that is later closed for the two cohorts of students who enrolled in and graduated from the to-be-closed middle schools three and four years prior to the closure announcement and phase-out process in each of the three years of closure (n=5,554), compared to two future cohorts of control students who would have attended the closed schools had they not closed and who instead attend the reconstituted schools that replace the closed schools or other middle schools (n=5,554). My analyses do not include students attending schools when closure is announced and during the phase-out process, nor those who leave the schools during this transitional time period.

					Schools cease admitting 6th graders and begin phase-out process. ⇩			**Schools close in 2006,** phase-out process ends. ⇩		
Student subsample	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Treatment: Attend to-be-closed school	4th grade 3rd grade	5th grade 4th grade	6th grade 5th grade	7th grade 6th grade	8th grade 7th grade	9th grade 8th grade	9th grade			
Attend during phase-out (transition)		3rd grade	4th grade 3rd grade	5th grade 4th grade	6th grade 5th grade	7th grade 6th grade	8th grade 7th grade	9th grade 8th grade	9th grade	
Control: Attend alternate receiving school				3rd grade	4th grade	5th grade	6th grade 5th grade	7th grade 6th grade	8th grade 7th grade	9th grade 8th grade

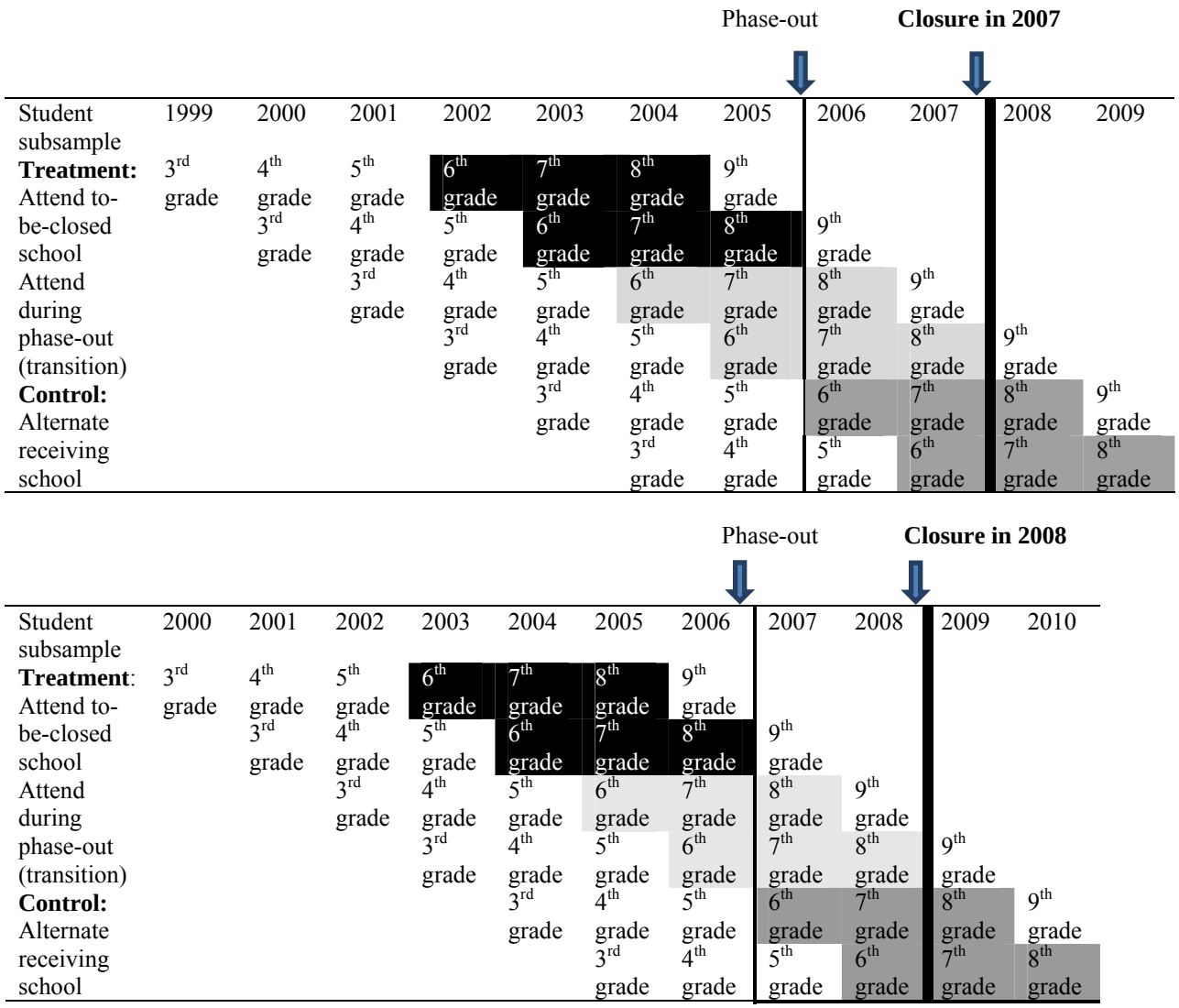


Figure 2-1. Student cohorts.

The key challenge in estimating the impact of school closure policies on student outcomes is identifying what would have happened in the absence of the reform. Ascribing causality requires disentangling a complex process involving multiple policy stages and student and school selection and enrollment processes, given that schools are phased out and phased in over several years. My identification strategy rests on a difference-in-differences matching approach (Heckman, Ichimura & Todd, 1997). I conceptualize the treatment group as sixth graders who

enrolled in a school that is later closed. A simple approach to examining the impact of attending a chronically underperforming school that is later closed would be to compare student test scores before and after the school is closed and reopened as a new school. The obvious limitation of this approach is the potential for selection bias. That is, it is reasonable to assume that students who attended schools selected for closure differ from those who attend the reconstituted schools. Indeed, opponents of school closure contend that the newly reconstituted schools do not in fact serve students in the community, and that they select only higher-performing students (ACNY & AALDEF, 2009; Plasencia & Agarwal, 2009). In addition to lower test scores, the New York City Department of Education might also have closed schools whose students and families possessed fewer social and political resources to oppose closure. Such resource disparities are likely related to student learning as well.

Matching. To address concerns about selection bias inherent in this topic, I match students who enrolled in chronically failing schools that were subsequently closed (the treatment group) to students likely to have enrolled in the closed schools (the counterfactual, or control group) had those schools been available. To approximate this counterfactual I employ propensity score matching methods to create a control group of students likely to have enrolled in the closed schools, but who did not have the opportunity.

Propensity score analysis involves matching cases based on their shared probability of receiving the treatment (here, enrolling in schools selected for closure) conditional given what we observe on the case prior to the treatment (the confounding covariates). This allows for the identification of a similar group of controls by matching treated cases and control cases on the probability of having received the treatment. Traditional linear regression techniques that condition on specific covariates ignore the fact that there may be little overlap between the

treatment and control groups—that is, the two groups may be too different along particular dimensions to allow for valid comparisons. Propensity score matching restricts these comparisons by creating a more balanced sample of matched groups based on particular characteristics, a method that does not require extrapolation across areas with minimal data (Gelman & Hill, 2007; Rosenbaum & Rubin, 1983).

Research on matching suggests that if the treatment is assigned at the school level, matching should be conducted at the school level to avoid biased results (Stuart, 2007). However, I match on a student-level measure of treatment, using three dummy-coded measures of whether the student enrolled in sixth grade in a school that was closed in 2006 (yes=1, no=0), 2007 (yes=1, no=0) and 2008 (yes=1, no=0). I do so because of the nature of school closure policies in New York City, in which the school that is treated ceases to exist, and the resulting counterfactual or control condition is not experienced by a school as a whole, but rather by students who enroll in multiple alternate schools that contain both control and non-control students. Additionally, I focus on *enrolling* in rather than *exclusively attending* a to-be-closed school, as I aim to estimate the impact of attending a to-be-closed school for students who stay the full three years (grades 6-8) as well as for those who leave after a shorter period. It seems reasonable that enrolling in a chronically underperforming school would increase (or possibly reflect) student mobility. By including "stayers" and "leavers" in my models, my results are inclusive of any potential effects of mobility on achievement. Therefore, my approach is an estimation of an intent-to-treat (ITT) effect for students enrolled in the schools at baseline.

I conduct matching separately for each year of closure, identifying three control groups of students who enrolled in schools that closed in 2006, 2007 and 2008 in sixth grade three and four years prior to the closure announcement. By conducting separate matching for each year of

closure, I allow for the possibility that enrollment patterns may differ across years. I perform matching using the "MatchIt" package in R (Ho, Imai, King & Stuart, 2011). In the first step, I estimate the probability of enrolling in a middle school that is later closed using a logistic regression. I account for a number of student-level covariates related to both treatment assignment and the outcome in order to identify the most plausible control group (Heckman, Ichimura & Todd, 1997; Ho, Imai, King, & Stuart, 2007). Models include students' state test scores in English and mathematics and rates of absences and tardiness for grades three, four and five; free and reduced-price lunch eligibility; whether the student ever received special education services; whether a language other than English was the primary home language; student mobility in elementary school; whether the student was old-for-grade at the beginning of sixth grade; and race/ethnicity to compute the probability each student would have enrolled in a chronically underperforming school later selected for closure. The estimate from this first step is the propensity score.

In the second step I use one-to-one nearest-neighbor matching without replacement and common support to restrict the sample to cases that do not have propensity scores below or above the range of propensity scores for the treatment cases (Rubin, 1973). Nearest-neighbor matching performs particularly well with datasets that include many more control units than treatment units, and if many of the controls are quite different than the treated individuals, which is the case with these data (Ho et al., 2007; Stuart, 2010). Because I discard unmatched control students, my estimates are the effect of the treatment on the treated.

I use a particularly restrictive approach to matching students by requiring an exact match on the students' elementary school, selecting the student with the closest propensity score within the same elementary school. Given the strong elementary-to-middle school attendance pipelines

in New York City, performing an exact match on elementary schools provides a considerably more plausible matching estimate than matching solely on student characteristics. Additionally, this approach implicitly controls for any unobserved differences between the treatment and control units that are related to elementary school choice, and that are likely associated both with family and student decisions to enroll in a chronically underperforming middle school, such as family support for learning at home. I also examine a few other more restrictive matching approaches. I restrain the matching using an exact match and second, using Mahalanobis metric matching within a caliper of 0.25-.05 SD to limit selection of matches to those with the smallest Mahalanobis distance within a small range of propensity scores. I do so for select important measures, test scores and attendance rates (Rubin & Thomas, 2000), but with these more restrictive approaches too many of the treated students (more than 300) remain unmatched, compromising the external generalizability of the study. Therefore, in order to retain the full sample of treated students, I relax the exact and Mahalanobis distance matching constraints. I also explore oversampling to reduce the variance by using one to two and one to three nearest neighbor matching (Smith, 1997) but the oversampling resulted in poorer balance than when using one to one nearest neighbor matching.

I conduct the matching for each of the three years of closure and for each of the five multiply imputed datasets. Matching using imputed data tends to result in less bias than matching using complete cases, and is equivalent to matching on missing data patterns (Graham 2008; Harder, Stuart & Anthony, 2010). To estimate the propensity scores, I use different logit models for each of the three years of closure cohorts to allow for optimal balance. I use the same logit models, however, across each of the five multiply-imputed datasets.

I use an iterative approach to matching, in which I specify the logit regression using the pre-treatment covariates, assess balance, and add interactions and higher-order polynomials for unbalanced measures to the logit model in order to address lack of balance. To take into account both means and distributions of the covariates, I compare the standardized difference in means--the means divided by the standard deviation of the treatment group--between the matched samples (Stuart, 2009). Research suggests that sufficient balance can be reached with an absolute standardized distance in means of no more .25 between the matched samples (Rubin, 2001; Stuart, 2009). Others suggest a more stringent threshold of .1 standardized differences in means (Austin, 2011; Normand et al., 2001). My matched samples meet the .25 threshold in all cases, and the .1 standardized difference threshold with a few exceptions for the 2006 and 2007 cohorts. Figures 2-2, 2-3 and 2-4 display the absolute standardized differences by student characteristic for each of the five imputed datasets by cohort. These differences are within-school differences because of the exact match on school attended in fifth grade. For the matched cohort of students who attended schools closed in 2006, absolute standardized differences in means are slightly greater than .1 for the second-level category within the five-level categorical measures of absences in fifth grade and tardies in third and fourth grade. For the matched cohort of students who attended schools closed in 2007, absolute standardized differences in means are between .1 and .2 for ever having an IEP, free/reduced lunch eligibility, home language, the second-level category within the five-level categorical measures of tardies in third and fifth grade and the second- and fourth-level category within the five-level categorical measure of tardies in fourth grade.

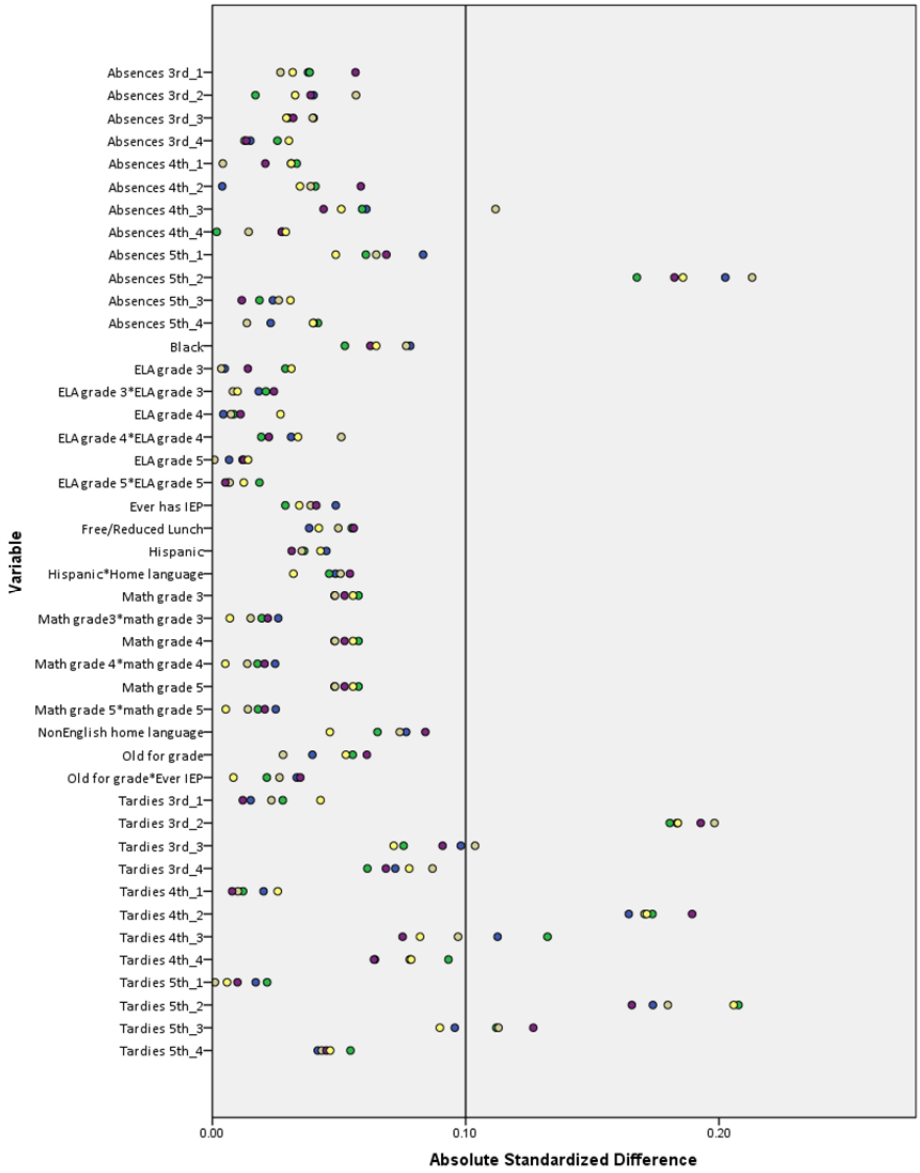


Figure 2-2. 2006 Cohort: Absolute standardized differences between matched sample means.

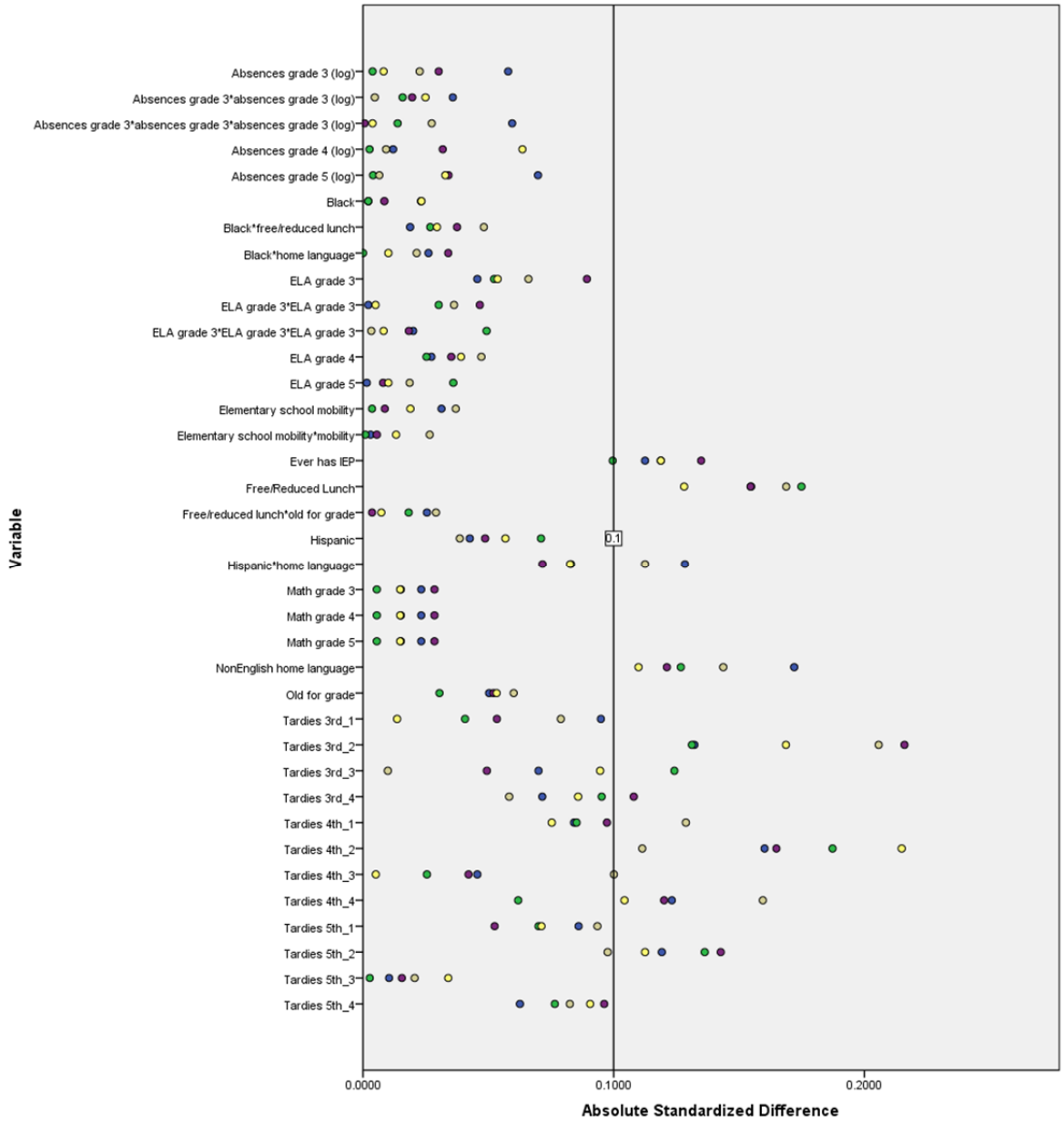


Figure 2-3. 2007 Cohort: Absolute standardized differences between matched sample means.

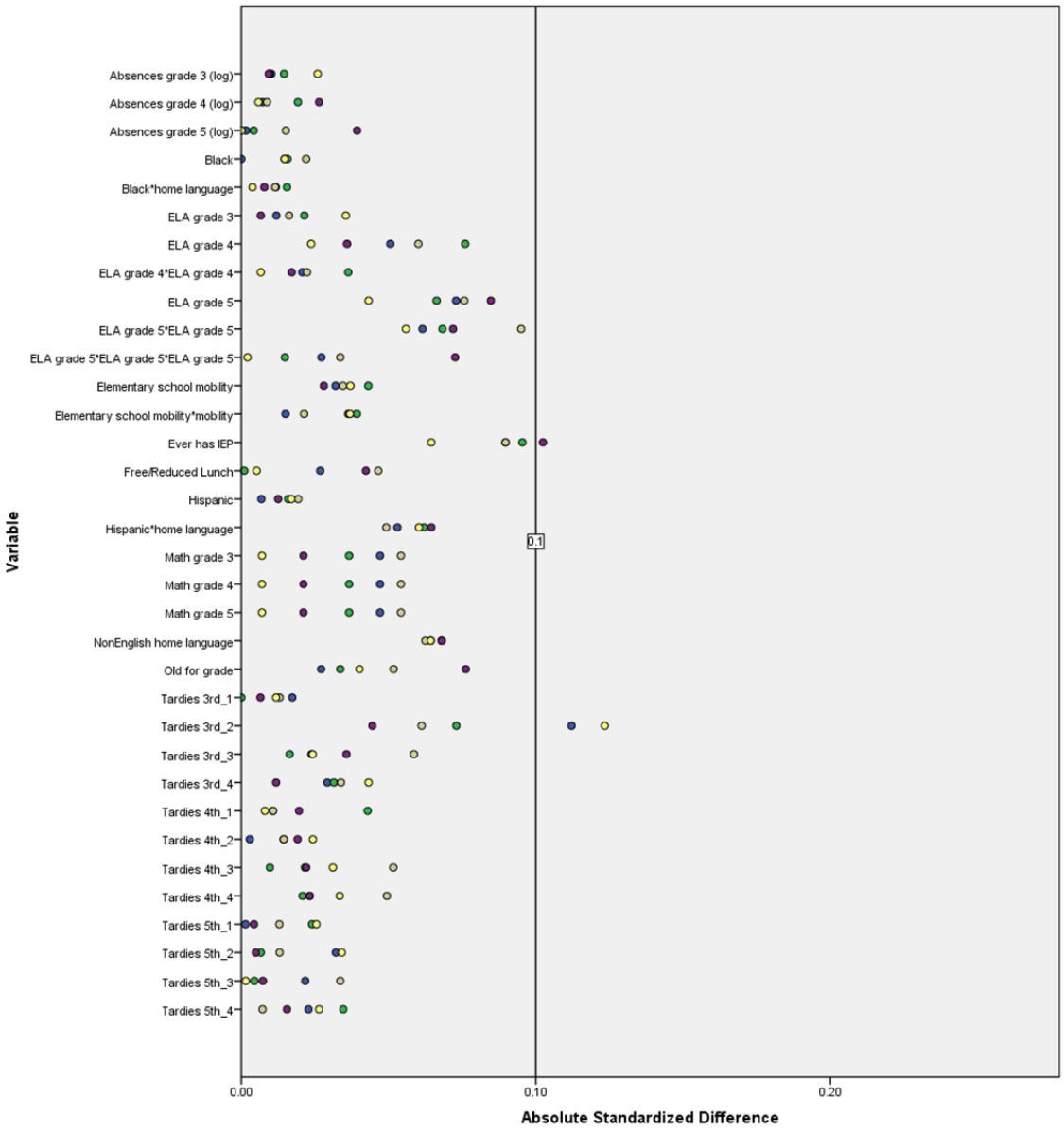


Figure 2-4. 2008 Cohort: Absolute standardized differences between matched sample means.

Modeling impact. Once I create a set of students matched to the treatment groups, I estimate student learning as a function of enrolling in a school that was later closed. For these analyses, I use a difference-in-differences approach within a multi-level framework (Heckman, Ichimura & Todd, 1998; Raudenbush & Bryk, 2002). Although it seems reasonable to assume that treatment and control groups are quite similar after matching via the propensity score

method, additional methods to account for the unmeasured differences in the students across these schools seem warranted, and research suggests that using both matching and covariate adjustments are more effective in addressing selection bias than matching alone (Rubin, 1979; Rubin & Thomas, 2000). Thus, I estimate the effect of the treatment using multi-level cross-classified models within a difference-in-differences framework in which student test scores from grades 5-8 are nested in students who are nested in schools, conditioning on student observed characteristics for a more robust estimate of the treatment effect (Hill, 2008; Rubin & Thomas, 2000; van der Laan & Robins, 2003). Multi-level approaches provide more reliable school-level estimates and standard errors than standard OLS techniques (Raudenbush & Bryk, 2002; Singer & Willett, 2003; Snijders & Bosker, 1999). The cross-classified nature of the models allows me to account for student mobility from school to school, as students can be nested in multiple schools as they progress from elementary school through middle school (Grady & Beretvas, 2010).

My difference-in-differences model compares rates of changes in test scores within the control group to rates of changes in test scores within the treatment group across time periods, pre- versus post-treatment, as students move from grade five (pretreatment or initial status) through eight. The impact of school closure on student learning is the difference between the treatment group's change in test scores and the control group's change. As indicated in equation 1, the basic difference-in-difference model can be written as:

$$\tau = (\overline{Y}_1^{t_1} - \overline{Y}_1^{t_0}) - (\overline{Y}_0^{t_1} - \overline{Y}_0^{t_0}) \qquad (1)$$

This approach requires one specific assumption to support interpretation of the estimate as a causal effect: that the change in mean test scores that the control group experiences over time reflects the same change that the treatment group would have experienced had they not been

exposed to the treatment. The matching conducted prior to the difference-in-difference analysis makes this assumption more plausible. Moreover, conditioning on student pre-treatment background characteristics when estimating the impact also makes this assumption more reasonable.

In a multi-level regression framework, the difference-in-differences estimation can be rewritten as:

Level 1 (repeated-measures):

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}(6th\ grade)_{tij} + \pi_{2ij}(7th\ grade)_{tij} + \pi_{3ij}(8th\ grade)_{tij} + e_{tij}$$

Level 2 (individuals): (2)

$$\pi_{0ij} = \beta_{00j} + \beta_{01ij}\,(CLOSED_{ij}) + \beta_{02ij}\,(X_{ij} - \bar{X}) \ldots + r_{00ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11ij}\,(CLOSED_{ij}) + \beta_{12ij}\,(X_{ij} - \bar{X}) \ldots + r_{10ij}$$

$$\pi_{2ij} = \beta_{20j} + \beta_{21ij}\,(CLOSED_{ij}) + \beta_{22ij}\,(X_{ij} - \bar{X}) \ldots + r_{20ij}$$

$$\pi_{3ij} = \beta_{30j} + \beta_{31ij}\,(CLOSED_{ij}) + \beta_{32ij}\,(X_{ij} - \bar{X}) \ldots + r_{30ij}$$

Level-3 (schools):

$$\beta_{00j} = \gamma_{00} + u_{0j}$$

$$\beta_{10j} = \gamma_{10} + u_{10j}$$

$$\beta_{20j} = \gamma_{20} + u_{20j}$$

$$\beta_{30j} = \gamma_{30} + u_{30j}$$

The level-1 model estimates the outcome at a given time point (grade), where *Ytij* is the achievement outcome or rate of absences of child *i* in school *j* at time *t*. π_{0ij} is the predicted value for individual *i* in school *j* when time (grade, in my model) equals 0, or fifth grade. To account for nonlinear changes in student test scores and absences, I use separate indicators of exposure to

each grade. These three measures of school exposure—sixth grade, seventh grade and eighth grade—permit the modeling of four distinct parameters: (1) *initial status* (π_{0ij}), or children's achievement or absences at fifth grade (literally, predicted achievement or absences with exposure to zero years of sixth grade, zero years of seventh grade, and zero years of eighth grade). The three remaining parameters (π_{1ij} , π_{2ij} and π_{3ij}) are linear attendance and learning rates or slopes over: (2) *sixth grade*; (3) *seventh grade*; and (4) *eighth grade*. I allow these rates to vary within students, and examine the impact of closure separately by the three years of closure. Because the student test score outcomes are standardized within subject, grade and school year, rates of learning here are relative to other students, the measure of 'learning' is not an indicator of absolute growth, but rather indicates whether the student's test scores moved up or down *relative* to the distribution of all student test scores within that subject, grade and school year.

The level-2 model estimates student learning as a function of the treatment (enrolling in a closed schools) and student demographic characteristics (X_{ij}), which are grand-mean centered. By interacting grade and whether a student enrolled in a school that was slated for closure, I am able to estimate whether students who enrolled in the schools to be closed experience differential rates of change between fifth grade and eighth grade—that is, whether the slope of their test score trajectories differ having enrolled in later-closed schools compared to enrolling in alternate receiving schools. At the school level (Level 3), γ_{00} is the average learning in the population and u_{0j} is the error term associated with school *j*. Although they account for the clustered nature of the data at the school level, these level 3 models are unconditional in that they do not adjust the treatment estimates for school-level characteristics. This is necessary for the simple fact that the characteristics of each school vary across the multiple cohorts it serves over time.

I implement these models using the "lme4" package available with R (Bates, Maechler & Dai, 2009). Additionally, I conduct these analyses across each of the five imputed datasets, and combined the results based on Rubin's (1987) approach using the "mitools" package in R (Lumley, 2010).

Fixed effects. One potential bias in the difference-in-differences analysis may stem from unobserved student or family characteristics that are related both to change in student outcomes over time and enrollment in a chronically-underperforming middle school. As a robustness check on these initial results, I include a second approach that measures the impact of school closures using an individual fixed-effects model that implicitly holds constant both observed and unobserved time-invariant student characteristics by comparing student test-score and attendance trajectories within individuals before and during middle school. In this approach, students serve as their own counterfactuals to what would have happened had they not attended a to-be-closed school. Changes in student outcomes can be interpreted as the change relative to a student's outcomes before or after he or she switched schools. An individual fixed-effects model improves the plausibility of the assumption that the assignment to the treatment—attending a chronically underperforming school—is ignorable because *both* observed and unobserved characteristics that are stable over time are controlled for (Gangl, 2009; Gelman and Hill, 2008; Winship & Morgan, 1999).[2]

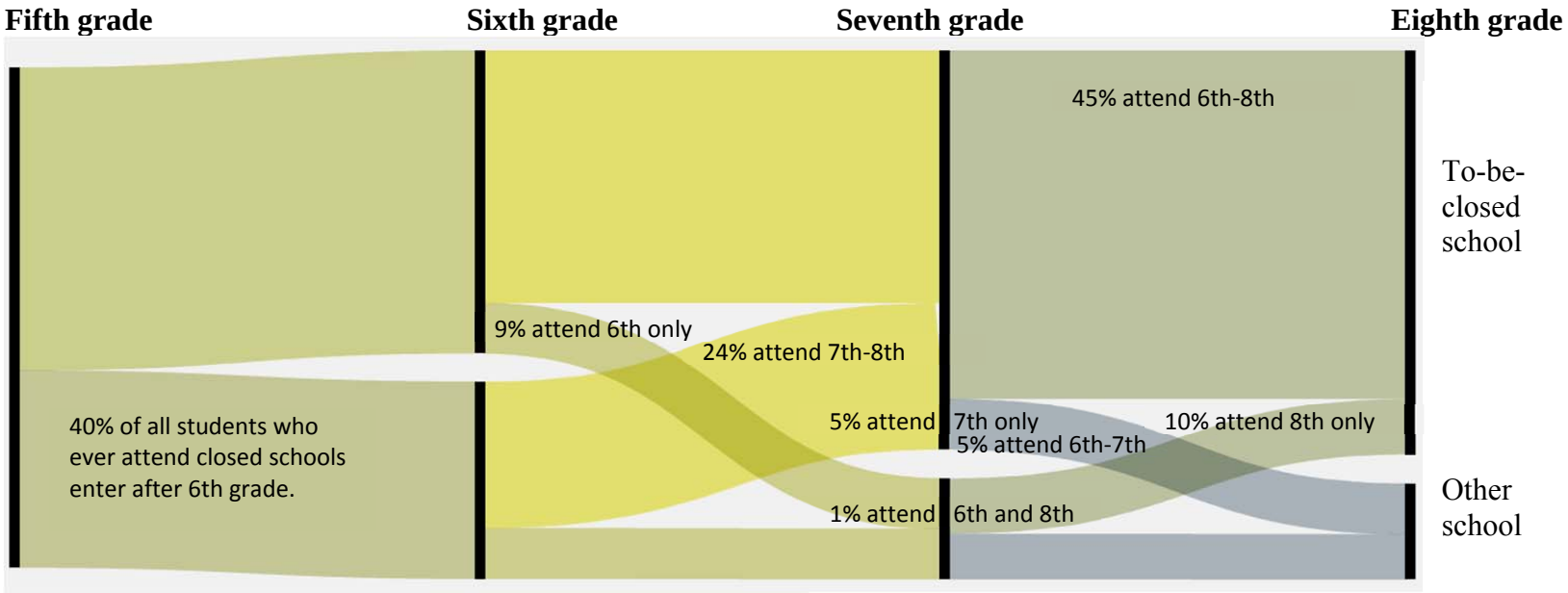
This second analysis answers a slightly different question for a slightly different student population. Although both models estimate the average effect of the treatment on the treated, my first analysis focused on the impact of closure for a future cohort of students who would have attended a chronically failing school had it not closed. Rather than focusing on the effects of

[2] Note that the original difference-in-differences approach is a version of a fixed-effect model using a student-level event (Gangl, 2009), enrolling in a to-be-closed school, rather than the incident-level fixed effects approach described above.

school closure *after* a school has closed, this analysis focused on the impact of attending a chronically failing school compared to the counterfactual of attending any other *current* alternative receiving school. The sample in this analysis is limited to students who were in sixth grade three and four years prior to the closure announcement and who ever enrolled in a school closed in 2006, 2007 or 2008 (n=9,591 students). Therefore, the impact is not necessarily an indication of the impact of fully closing a school, but rather, of attending a to-be-closed school compared to its contemporary alternatives. If a whole school were closed it seems likely that some alternate receiving schools would be flooded with a host of high-need students who would have attended the closed school had it not closed. This alternate model does not account for these kinds of consequences. Rather, this analysis exploits the high-levels of student mobility in the treatment population to examine differences in learning and attendance trajectories as students switch in and out of treatment schools. Additionally, the fixed-effect analysis also focuses on a slightly broader treatment of attending a to-be-closed school at any point during middle school, while the first analysis focused on the impact of enrolling in 6th grade in a to-be-closed school.

Although an individual fixed-effects approach may allow for more believable assumptions about ignorability of treatment assignment, it also sacrifices some external validity and limits the generalization of findings to a smaller population of students. The impact of closure is identified only when students change from control to treatment status or vice versa. Therefore, the results of this analysis only apply to those students who switched from a to-be-closed school to a school that was not selected for closure or vice versa at some point during fifth through eighth grade. These mobile students are not a particularly unusual population, however. Figure 2-2 below illustrates the high levels of student mobility in the closed schools in this sample. Student mobility patterns in schools that were later closed suggest that of the students

who ever enrolled in one of these schools during the treatment years, more than half (55%) switched schools during sixth through eighth grade. Just under half (45%) attended all three years, one-quarter of students attended during seventh through eighth grade, and a final one-quarter of students attended for just one year (sixth, seventh or eighth grade only).



(Developed using Fineo, DensityDesign Research Lab - Politecnico di Milano)

Figure 2-5. Student mobility in to-be-closed schools.

As is typical of more mobile students, the students who stay at a to-be-closed school before closure is announced for all three years are slightly more advantaged than those who move during middle school; the movers' fifth grade ELA and mathematics test score are approximately 0.2 SD lower than their less mobile peers, on average. These more mobile students also have slightly higher rates of absences and tardiness in fifth grade, are more likely to have English as a home language, and are slightly more likely to be old for their grade upon entering sixth grade. Thus, the estimates from these fixed-effects models are for a slightly more disadvantaged population of students.

Because of the nested nature of the data, for the fixed-effects analyses I use a multi-level approach that is functionally equivalent (but statistically more robust) to traditional OLS fixed-effect approaches. Specifically, I use adaptive centering with random effects (Raudenbush, 2009) in which I group-mean center the incidence-level treatment measures, a dummy-measure of whether the student was enrolled that year at a to-be-closed middle school (coded 1 if attending a to-be-closed school, 0 otherwise). For these analyses, I use a two-level growth-curve framework in which student outcomes—ELA and mathematics state test scores and logged absences—from grades 5-8 are nested in students. Adaptive centering offers an improvement over traditional fixed-effects models in two ways. First, it allows for modeling heterogeneous treatment effects across students by using random effects (Raudenbush, 2009). Additionally, accounting for the clustered nature of the data provides more efficient and reliable estimates and standard errors compared to standard OLS regression fixed-effect analyses (Raudenbush, 2009; Raudenbush & Bryk, 2002; Singer & Willett, 2003; Snijders & Bosker, 1999).

By interacting a students' exposure in prior years to a chronically underperforming school with the treatment indicator, I am able to estimate whether students who attend a to-be-closed

school for multiple years experience differential rates of learning—for example, whether additional years at a chronically underperforming school compounds the impact of the treatment. This incident-level variable is coded 0 for all years prior to entering a to-be-closed school, and increases by one each year after the student was enrolled in a to-be-closed school. Finally, to help satisfy ignorability of treatment assignment, I also control for grade to account for the fact that students are more likely to move schools and to be absent from school in later years. I also include cohort fixed effects to account for differences across years. I conduct the fixed effects analyses across each of the five imputed datasets using HLM 6.0 (Raudenbush & Bryk, 2002).

Chapter 3: Characteristics of Closed and Reconstituted Schools

This chapter focuses on describing implementation of school closure and reconstitution in New York City. I examine the characteristics of the schools prior to closure and explore the characteristics of the schools that replace the closed schools. Unsurprisingly, I find that schools selected for closure have significantly lower school-average state test scores and lower attendance rates compared to other middle schools, in almost all cases for several years prior to selection for closure. Furthermore, schools selected for closure enroll students with significantly lower levels of achievement prior to middle school compared to students in other schools, lending credence to the claim that these schools evidence lower achievement in part because they serve high-need students who are already at a disadvantage prior to entering the middle school. I also find that the reconstituted schools that replace the closed schools, while similar demographically, tend to enroll students who are less disadvantaged academically, although this trend is true only for schools closed and reconstituted in 2004 and later.

This descriptive chapter uses all data on closed schools from 1999 through 2008, with a focus specifically on middle schools, which I define as schools whose lowest grade is between four and seven, and whose highest grade is between five and nine. I present four sets of analyses. I start with a description of the number of middle schools that have closed in the District, and describe the academic and demographic characteristics of schools that were closed, comparing characteristics of to-be-closed schools to other middle schools. I also examine whether the characteristics of schools selected for closure change across time as the school closure policy matures.

Second, I examine the extent to which these schools are chronically underperforming by describing the academic characteristics of schools in the years prior to closure. New York City has only recently clarified its policies and processes to identify school closures, partially in response to changes in state legislation calling for increased transparency. District administrators characterize decisions to close schools as based primarily on a sustained lack of academic achievement, which to them suggests that the staff in these schools lack the capacity or will to improve. However, the district has historically faced and continues to face contentions that the decisions to close schools are biased and based on political favoritism rather than more objective rationales, and further, that the reform unfairly targets schools that enroll concentrations of minority students. I examine implementation of closure and the characteristics of closed schools over time with these charges in mind.

Third, I describe the academic characteristics of students prior to enrolling in the to-be-closed schools with the purpose of determining the extent to which the students who enroll in the chronically low-performing schools enter into the school already at a disadvantage. Finally, I describe school reconstitution. The rationale for school closure rests on the assumption that the reconstituted schools that replace the closed schools will provide more effective academic support and result in higher student academic achievement. Opponents to school closure argue that the reconstituted schools are not more effective and, moreover, that any evidence of higher achievement stems from the fact that the new schools do not serve high-need students, such as the lowest-achieving students or English-language learners. To provide insight into these claims, I examine the characteristics of the reconstituted schools, comparing these new schools and the students who attend them to the characteristics of the closed middle schools and their students.

Schools Closed in New York City from 1998-2008

As displayed in Figure 3-1, the New York City DOE closed 66 middle and junior high schools between 1997 and 2008. The majority of the closed schools were located in Brooklyn and the Bronx, followed by Manhattan. No middle schools in Staten Island were selected for closure during this time period. Additionally, school closure activities did not appear to vary much across administrations; the Bloomberg/Klein administration closed a mean of six middle schools per year compared to five middle school closures per year in the prior administrations.

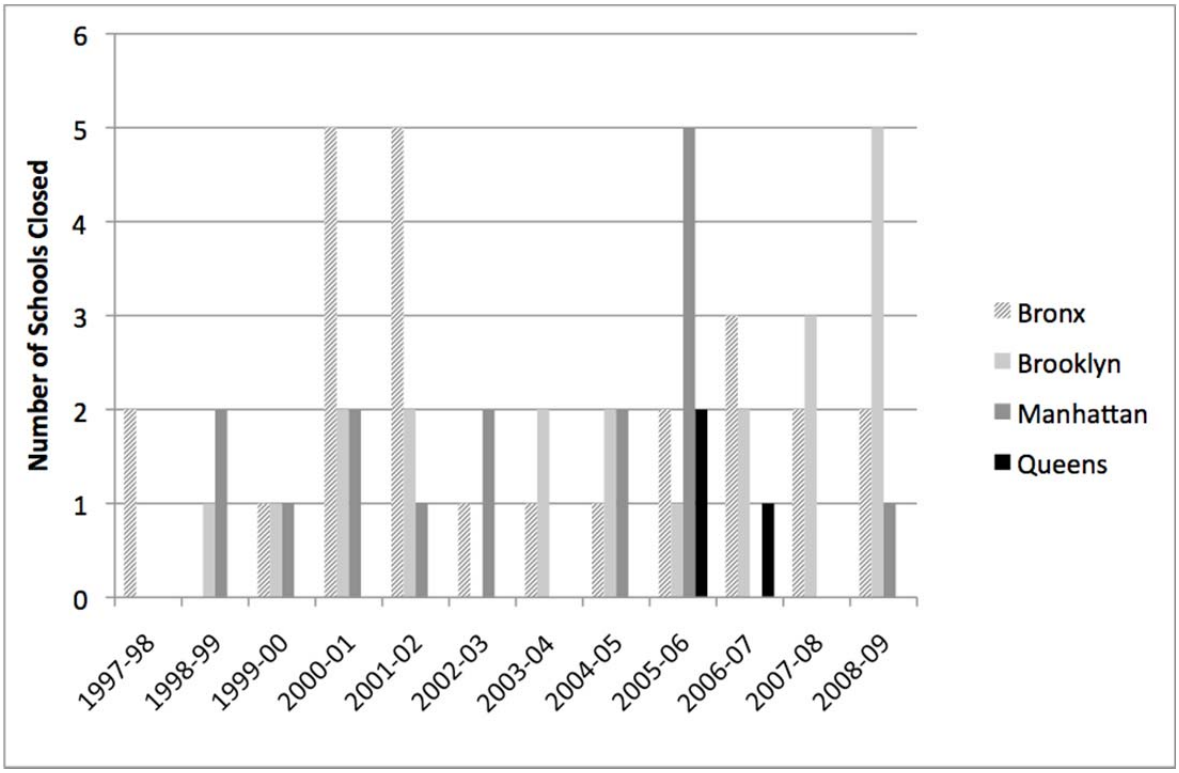


Figure 3-1. Middle school closures by borough.

Characteristics of closed schools

Table 3-1 displays the academic and demographic characteristics of closed middle schools over time, compared to other regular middle schools.[3] I examined the characteristics of the closed schools three years prior to closure compared to all others middle schools that year. Three years prior to closure is the approximate time of the announcement of closure for those schools that phase-out over a few years. My results suggest school-average test scores were significantly lower than those of other regular middle schools: closed schools had school-average English language arts test scores close to one-half standard deviation below other middle schools in most years.[4] A similarly-sized gap was evident in math scores between closed and other middle schools, although for a few years—2001, 2002 and 2003—the gap was closer to one-third standard deviation. Additionally, school-average student absences and tardies were also significantly higher at closed schools: school-average absences for closed schools were approximately two days higher per semester than other schools, and closer to three days more after 2004. Similarly, school-average tardies were, with the exception of one semester, higher by two to six incidences per semester in closed schools compared to other middle schools. Although these school-level differences were substantively quite large, the necessarily small school sample sizes preclude the ability to identify statistically significant differences. Rather, when possible, I report differences in effect-size units or other meaningful metrics.

[3] Not included in this table are schools that serve middle-grades students but with wider grade spans, such as K-12, K-8, 6-12 schools. I also did not include in the middle school comparison group any charter schools or selective schools.

[4] I standardized the ELA and mathematics exam scores by computing z-scores within each grade, within each year. At the student level, the mean average score is 0. To present school-average test scores, I aggregated the individual scores, already standardized within grade and year, by school. I chose not to aggregate the raw scores for all students in the school and then standardize the school-average state test scores because of the differences in the state test and the metric across grades, which are not necessarily comparable. Thus, the aggregated, school-average standardized (z-scored) test scores are not centered at zero, and tend to be slightly lower than zero. This discrepancy is likely due to the slight positive association between school size and achievement (with the exception of the negative associations in 2004 and 2005), and the increase in the number of small schools over the period.

Table 3-1

School Academic Characteristics for Closed Schools Three Years prior to Closure (Estimated Time of Announcement)

School year	1998-99		1999-2000		2000-01		2001-02		2002-3	
Characteristics	Closed: last year is 2000-1 (n=9)	All other MS (n=176)	Closed: last year is 2001-2 (n=8)	All other MS (n=198)	Closed: last year is 2002-3 (n=3)	All other MS (n=203)	Closed: last year is 2003-4 (n=3)[b]	All other MS (n=202)	Closed: last year is 2004-5 (n=4)[c]	All other MS (n=200)
School-average ELA scores (mean)[a]	-.631	-.051***	-.434	-.060†	-.566	-.039	-.187	-.027	-.405	-.029
School ELA standard deviation (mean)	.907	.888	.954	.885	.837	.882	.834	.881	.866	.874
School-average math scores (mean)[a]	-.630	-.038	-.427	-.065†	-.513	-.040	-.150	-.039	-.391	-.024
School math standard deviation (mean)	.924	.889**	.955	.877	.844	.876	.832	.870	.960	.880
School-average days absent (mean)										
Fall	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	8.77	6.5
Spring	n/a	n/a	n/a	n/a	n/a	n/a	10.3	8.4	10.5	8.3
School-average days tardy										
Fall	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	7.3	11.1
Spring	n/a	n/a	n/a	n/a	n/a	n/a	9.3	7.7	14.5	9.1†

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Measure is z-scored (*M*=0; *SD*=1) at the student level.

[b] Student test scores, attendance, mobility, language and age data for JHS 142 Stranahan are not available for this school year.

[c] Data for JHS 17 LaSalle School are not available.

School year	2003-4		2004-5		2005-6		2006-7	
Characteristics	Closed: last year is 2005-6 (n=10)	All other MS (n=201)	Closed: last year is 2006-7 (n=6)	All other MS (n=245)	Closed: last year is 2007-8 (n=5)	All other MS (n=262)	Closed: last year is 2008-9 (n=8)	All other MS (n=267)
School-average ELA scores (mean) [a]	-.422	-.025*	-.477	-.049†	-.470	-.089	-.479	-.079*
School ELA standard deviation (mean)	.808	.868	.863	.835	.860	.848	.831	.840
School-average math scores (mean) [a]	-.341	-.017†	-.501	-.038*	-.561	-.098†	-.531	-.097*
School math standard deviation (mean)	.868	.874	.928	.831	.872	.835	.810	.817
School-average days absent (mean)								
Fall	8.4	6.5	8.4	6.0	9.5	6.7	8.9	6.0**
Spring	11.0	8.1	11.8	8.7	12.2	8.6	11.8	8.4*
School-average days tardy								
Fall	9.6	6.7*	8.4	5.7	9.1	6.5	13.2	6.9***
Spring	12.6	8.6**	11.6	8.0	14.7	8.9*	14.1	8.9*

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Measure is z-scored (*M*=0; *SD*=1) at the student level.

Figures 3-2 and 3-3 below display the school-average ELA test scores by year, for each closed school, to depict the tremendous variation in the school-average scores of closed schools. The figure suggests that, with a few exceptions, most schools' school-average ELA and math scores were consistently below the mean of other middle schools not selected for closure (marked by the heavy black line). However, the spread of school-average achievement for closed schools varied by close to one-half standard deviation; the bulk of closed schools had school-average test scores between -0.25 and -0.75 standard deviations. Only a few schools started out with high school-average test scores and descended rapidly in the years prior to closure, and in a few cases, school-average ELA and mathematics scores for the school improved prior to closure. (I discuss this trend later in the dissertation.)

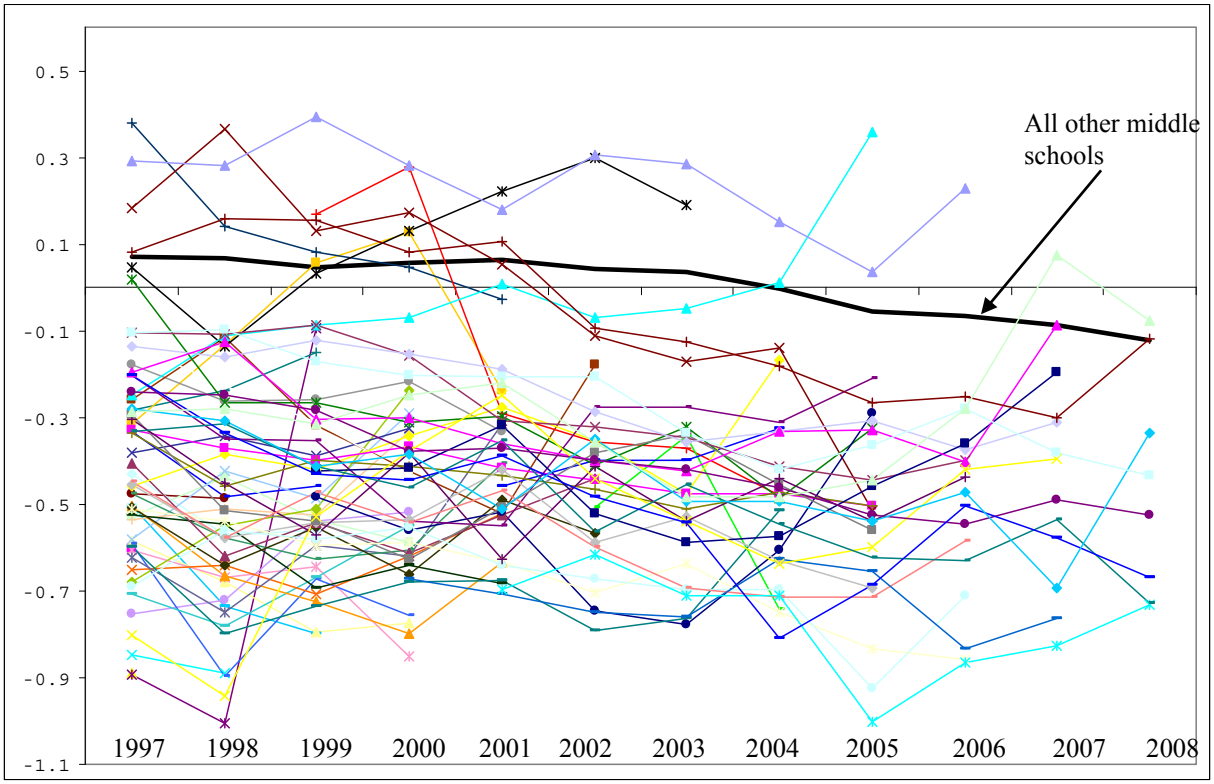


Figure 3-2. School-average ELA scores of closed schools, 1998-2008. [a]

[a] Measure is z-scored (M=0; SD=1) at the student level.

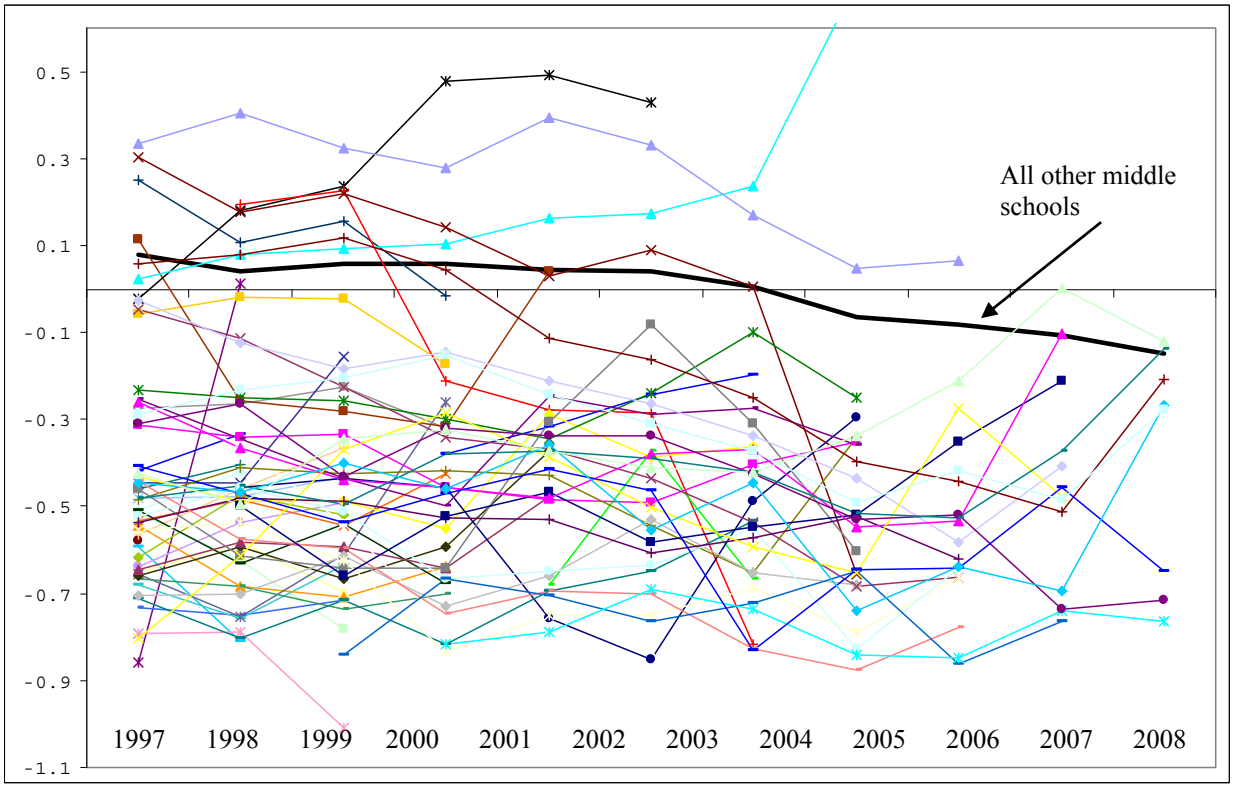


*Figure 3-3.*School-average mathematics scores of closed schools, 1998-2008. [a]
[a] Measure is z-scored (*M*=0; *SD*=1) at the student level.

Some observers have posited that vocal objections to closures have increased over the past few years because the closed schools are no longer clearly the lowest-performing schools—that is, that as the worst schools are closed, the schools closed in later rounds might be less clearly dysfunctional. Schools closed earlier—between 1998 and 2001—had somewhat lower school-average ELA scores compared to schools closed later in the time period of this study; however, the same pattern was not evident with math scores. Across both content areas, the smallest disparities in school-average achievement between closed schools and other schools, (<0.2 SD) occurred in 2001, and the largest gaps, approximately 0.6 standard deviations, evident in 1998. The mean standard deviations of school-average test scores were similar in both types

of middle schools, suggesting few differences in the within-school variation of student test scores.

I also examined the demographic characteristics of closed schools compared to other middle schools (see Table 3-2). Contrary to the perception of chronically underperforming schools as typically large, closed middle schools tended to enroll fewer students on average compared to other middle schools. Schools closed in later years were slightly larger, on average, compared to schools closed in earlier years.

However, fitting with the stereotype associated with failing urban schools, closed middle schools served fewer White and Asian students than other middle schools. In earlier years, closed schools tended to serve higher percentages of Hispanic students on average, compared to other schools, but in 2004 and later, this trend reversed, when middle schools selected for closure in these years began to serve fewer Hispanic students on average than other middle schools, and higher percentages of Black students. Regarding home language, no clear pattern emerged between closed schools and other schools, with differences appearing to vary from year to year. However, from 2004 on, closed schools tended to serve lower percentages of students whose home language is not English. And perhaps not surprisingly, middle schools selected for closure had slightly higher percentages of students eligible for free and reduced lunch than other types of middle schools. The percent of students eligible for free and reduced lunch was between five to twelve percentage points higher at closed schools than at other district schools. Closed schools also had more mobile students. On average, schools selected for closure had higher percentages of students who have moved more than once in elementary school compared to other middle schools.

Table 3-2

School Demographic Characteristics Three Years Prior to Closure (Estimated Time of Announcement)

School year	**1998-99**		**1999-2000**		**2000-01**		**2001-02**		**2002-3**	
Characteristics	Closed: last year is 2000-1 (n=9)	All other MS (n=176)	Closed: last year is 2001-2 (n=8)	All other MS (n=198)	Closed: last year is 2002-3 (n=3)	All other MS (n=203)	Closed: last year is 2003-4 (n=3)[a]	All other MS (n=202)	Closed: last year is 2004-5 (n=4)[b]	All other MS (n=200)
Enrollment (mean)	767	1041	712	966	476	986	685	963	421	1009
Racial composition (mean)										
Percent American Indian	0.3	0.2	0.3	0.3	0.2	0.2	0.5	0.3	0.7	0.3
Percent Asian	1.3	9.9	2.6	9.2	2.6	9.7	7.8	10.1	1.6	10.3
Percent Black	40.6	36.5	33.0	38.3	51.7	37.6	23.3	36.8	54.1	36.3
Percent Hispanic	57.1	37.4†	56.6	38.0	44.0	38.3	47.6	39.0	41.6	39.7
Percent White	0.8	16.0†	7.6	14.2	1.6	14.4	20.7	13.8	2.0	13.5
Percent eligible for free/reduced lunch (mean)	n/a	n/a	90.3	78.9	52.0	76.5	92.0	82.9	n/a	n/a
Percent old for grade (mean)	31.2	17.6**	16.3	12.2	22.0	11.7	15.0	12.2	18.5	12.4
Percent non-English household (mean)	8.7	5.1	12.2	7.1*	7.1	8.9	11.2	10.9	7.3	12.5
Percent with 2 or more moves in elementary school (mean)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	55.5	48.2

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Student test scores, attendance, mobility, language and age data for JHS 142 Stranahan are not available for this school year.

[b] Data for JHS 17 LaSalle School are not available.

School year	2003-4		2004-5		2005-6		2006-7	
Characteristics	Closed: last year is 2005-6 (n=10)	All other MS (n=201)	Closed: last year is 2006-7 (n=6)	All other MS (n=245)	Closed: last year is 2007-8 (n=5)	All other MS (n=262)	Closed: last year is 2008-9 (n=8)	All other MS (n=267)
Enrollment (mean)	818	961	878	790	589	706	644	648
Racial composition (mean)								
Percent American Indian	0.5	0.4	0.6	0.5	0.9	0.6	0.9	0.5
Percent Asian	4.2	10.5	5.9	9.3	1.7	9.6	1.6	10.3
Percent Black	48.1	35.5	53.3	37.3	61.4	37.1	62.4	36.5†
Percent Hispanic	45.5	40.3	37.1	42.1	33.4	42.4	32.1	42.4
Percent White	1.6	13.3	3.1	10.9	2.6	10.4	2.9	10.3
Percent eligible for free/reduced lunch (mean)	n/a	n/a	n/a	n/a	89.0	78.0	83.3	78.8
Percent old for grade (mean)	19.9	12.6**	21.7	13.4*	22.7	13.5*	20.4	13.6*
Percent non-English household (mean)	18.0	13.2	13.7	14.3	11.3	15.2	11.4	17.6
Percent with two or more moves in elementary school (mean)	55.1	49.4	57.1	50.0	61.0	50.6	56.7	50.1

†p<.10; *p<.05; **p<.01; ***p<.001

This initial analysis of school demographics and achievement suggests that in the approximate year closure was announced, to-be-closed schools were performing quite poorly. My results indicate that three years prior to closure, to-be-closed schools had substantially lower-performing students on average compared to other schools in the District. With a few exceptions, the gaps in school-average academic characteristics between closed schools and other middle schools appear consistent for 1998-2008, and varied between 0.37 and 0.58 SDs, with a smaller 0.16 SD difference in 2001. However, descriptions of the DOE policy for school closure suggest that the district administration reserves closure as a policy for schools that are *chronically* underperforming across multiple years. The next set of analyses test this assertion by examining the academic performance of the to-be-closed schools across several years prior to closure.

Closed Schools and Chronic Low-Performance

Tables 3-3 and 3-4 display school-average ELA and mathematics test scores and school enrollment for the years prior to closure, by the schools' last year of operation. In almost all years, schools slated to be closed evidence consistently low school-average test scores, generally between one-third to one-half standard deviation below average in both math and ELA several years prior to selection for closure. One exception to this pattern are the schools closed in 2003, which tend to have higher school-average test scores compared to schools closed during other years, close to one-fifth to one-quarter standard deviation below average in the years prior to the closure announcement. This exception to the overall trend appears to be driven by the closure of one school, Brooklyn Community School. Other schools closed this year are lower performing, consistent with the overall trend that suggests that closures were largely driven by accountability

purposes rather than the declining enrollment and fiscal considerations that drive closures in other parts of the country.

Table 3-3

Longitudinal Characteristics of Middle Schools by Year of Closure: School-Average ELA Test Scores

Last year of operation	School-average ELA scores by year											
	1997***	1998***	1999***	2000***	2001***	2002***	2003***	2004***	2005***	2006***	2007*	2008*
Never Closed	.070	.0670	.0448	.055	.065	.043	.035	-.001	-.056	-.068	-.087	-.122
1998	-.739	-.794	Closed									
1999	-.387	-.507	-.493	Closed								
2000	-.619	-.631	-.597	-.544	Closed							
2001	-.362	-.429	-.434	-.417	-.394	Closed						
2002	-.391	-.459	-.472	-.566	-.513	-.395	Closed					
2003	-.156	-.236	-.222	-.209	-.187	-.207	-.319	Closed				
2004	-.460	-.384	-.126	-.033	-.341	-.405	-.398	-.429	Closed			
2005	-.252	-.342	-.367	-.393	-.376	-.429	-.422	-.408	-.361	Closed		
2006	-.295	-.327	-.319	-.324	-.417	-.400	-.438	-.477	-.570	-.461	Closed	
2007	-.378	-.408	-.345	-.385	-.366	-.479	-.534	-.501	-.470	-.479	-.350	Closed
2008	-.195	-.204	-.267	-.291	-.331	-.384	-.446	-.514	-.556	-.479	-.466	-.452

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Measure is z-scored (*M*=0; *SD*=1) at the student level.

Table 3-4

Longitudinal Characteristics of Middle Schools by Year of Closure: School-Average Mathematics Test Scores

Last year of operation	School-average Math scores by year										
	1998***	1999***	2000***	2001***	2002***	2003***	2004***	2005***	2006***	2007*	2008†
Never Closed	.080	.041	.058	.058	.044	.039	.007	-.066	-.082	-.105	-.148
1998	-.742	Closed									
1999	-.494	-.493	Closed								
2000	-.630	-.606	-.627	Closed							
2001	-.374	-.427	-.404	-.402	Closed						
2002	-.397	-.478	-.513	-.516	-.270	Closed					
2003	-.246	-.208	-.153	-.150	-.154	-.207	Closed				
2004	-.433	-.147	-.131	-.382	-.391	-.322	-.511	Closed			
2005	-.315	-.366	-.355	-.408	-.373	-.341	-.318	-.314	Closed		
2006	-.295	-.293	-.353	-.472	-.433	-.467	-.501	-.609	-.550	Closed	
2007	-.363	-.401	-.498	-.416	-.450	-.498	-.514	-.561	-.521	-.393	Closed
2008	-.329	-.329	-.328	-.360	-.375	-.415	-.488	-.562	-.531	-.499	-.393

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Measure is z-scored (*M*=0; *SD*=1) at the student level.

Figure 3-4 displays mean differences between school-average ELA scores of closed schools and other middle schools by year of closure (represented by the zero line). In most cases, closed schools displayed school-average ELA scores that are consistently lower than those of other middle schools in the years prior to closure. For schools closed in all years but 2003, school-average ELA scores a few years prior to closure were between one-third to more than one-half standard deviation below other middle schools. School-average mathematics scores reflect similar trends across all years, with the exception of schools closed in 2003, when school-average mathematics gaps are closer to one-quarter standard deviation (see Figure3-5).

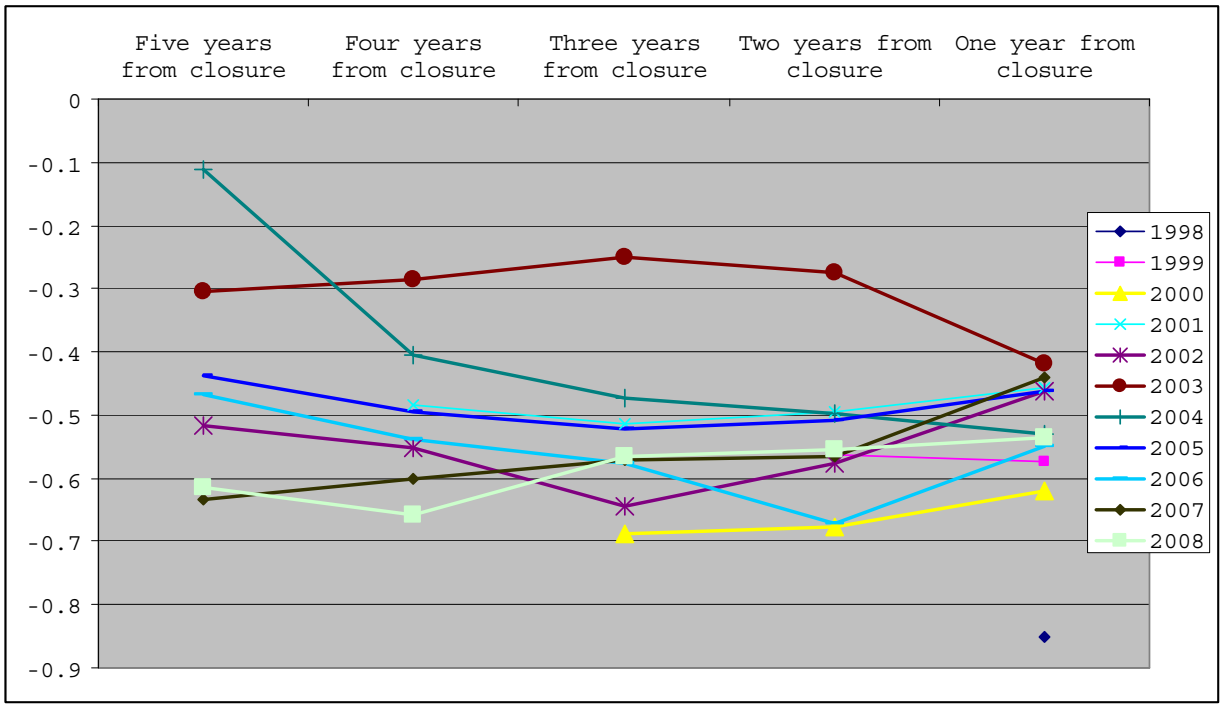


Figure 3-4. School-average ELA test scores: Mean difference from schools never closed.

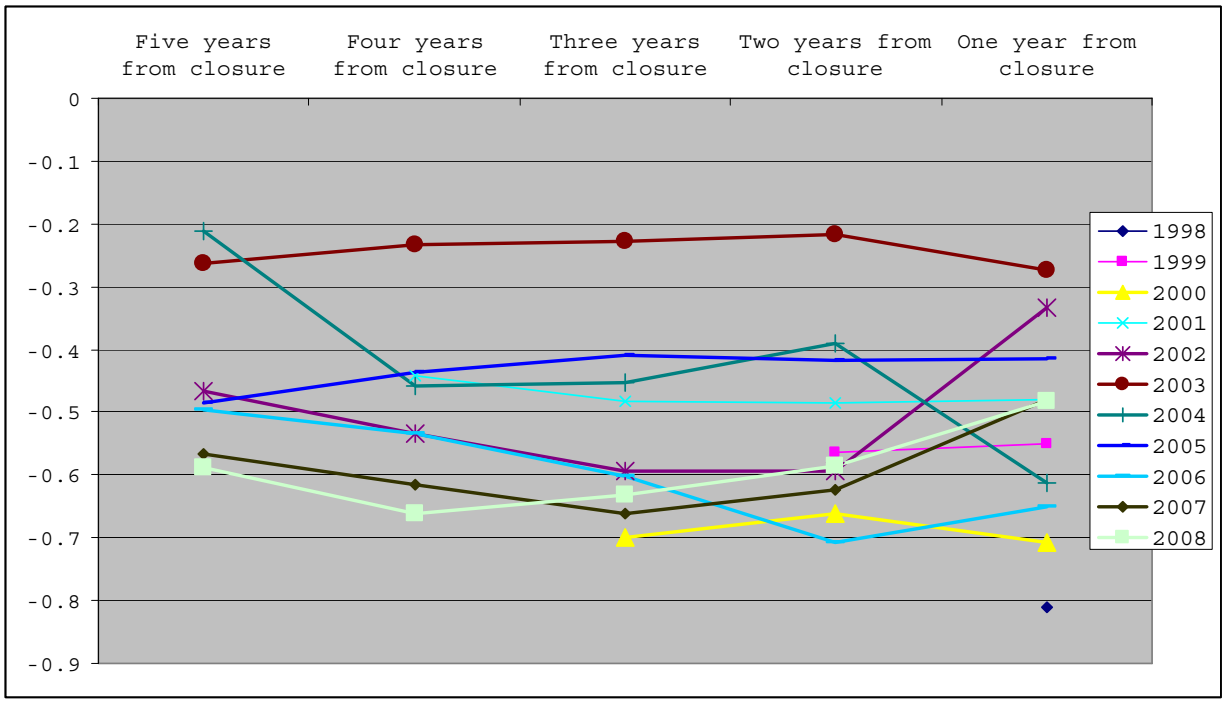


Figure 3-5. School-average mathematics test scores: Mean difference from schools never closed.

School-average test scores did not consistently descend in the years prior to closure. Figure 3-6 displays the mean growth (or drop) in school-average ELA and mathematics scores from the year prior; no clear pattern of growth or decline in school-average scores in the years prior to closure is apparent. Moreover, in most years, on average, school-average ELA and mathematic scores rose the year or two prior closure, which is likely the phase-out period for many schools. Similar trends were evident for mean differences in school-average mathematics test scores over time (see Figure 3-7). This growth in test scores during phase-out is somewhat counterintuitive in that one might expect more able students to be the first to abandon the school during phase-out. I discuss this unexpected pattern in more detail in the next chapter, which focuses on the process of school-phase out.

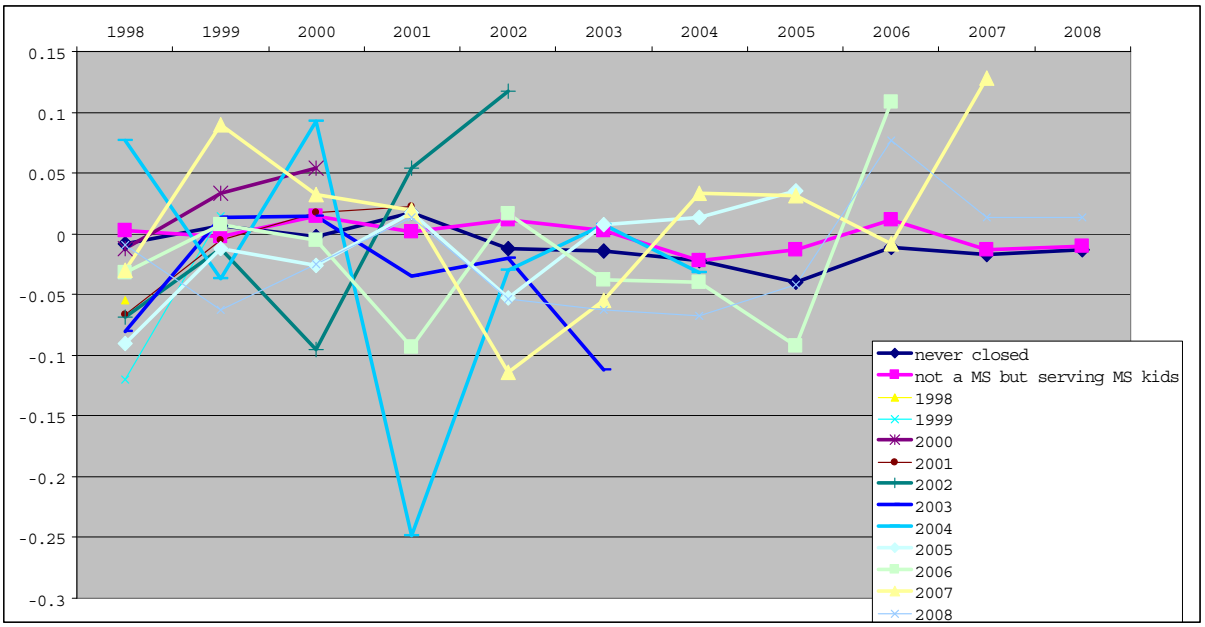


Figure 3-6. School-average ELA scores: Mean difference from prior year.

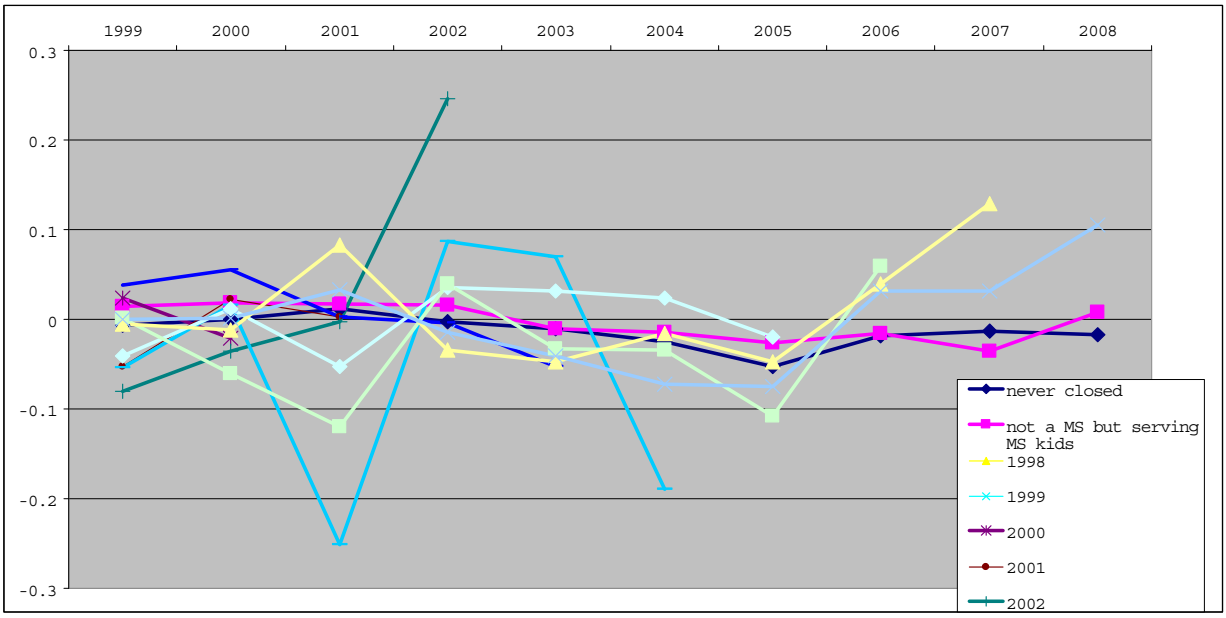


Figure 3-7. School-average Mathematics scores: Mean difference from prior year.

Differences in Enrollment Growth

The above findings suggest that with few exceptions, schools slated for closure tended to struggle with poor academic performance for several years prior to closure. However, in addition to academic achievement, New York City school administrators also cite declining

enrollment as a secondary indicator to identify schools for closure. Table 3-5 displays school enrollment across 1998-2008 by year of closure. Excluding the estimated phase-out years, for which school enrollment declined necessarily as grades were eliminated, my results suggest that in the years prior to closure the average enrollment remains to decline slightly, with the exception of schools closed in 2002, for which enrollment appears relatively constant. Note that average enrollment for other middle schools was also lower in later years of the study. However, unlike the closed schools, other middle schools did not appear to lose enrollment. Figure 3-8 displays the mean growth (or drop) in enrollment from the year prior, by year of closure (see also Table 3-6; approximate phase-out years are not displayed).

Table 3-5
Longitudinal Characteristics of Schools by Year of Closure: Enrollment

Last year of operation	Average Enrollment by year											
	1997***	1998***	1999*	2000***	2001***	2002	2003***	2004	2005†	2006*	2007**	2008***
Never Closed	1119	1121	1038	1035	1024	1044	997	810	718	660	603	598
1998	346	326	Closed									
1999	637	565	404	Closed								
2000	808	767	534	293	Closed							
2001	728	769	712	696	517	Closed						
2002	472	476	477	476	332	155	Closed					
2003	440	743	879	1159	685	690	348	Closed				
2004	1019	970	780	772	415	421	365	149	Closed			
2005	1072	1017	839	856	848	857	818	540	255	Closed		
2006	990	984	983	959	939	974	947	878	563	248	Closed	
2007	939	955	807	744	721	707	703	684	589	362	201	Closed
2008	785	843	874	886	830	895	882	834	752	644	394	184

†p<.10; *p<.05; **p<.01; ***p<.001

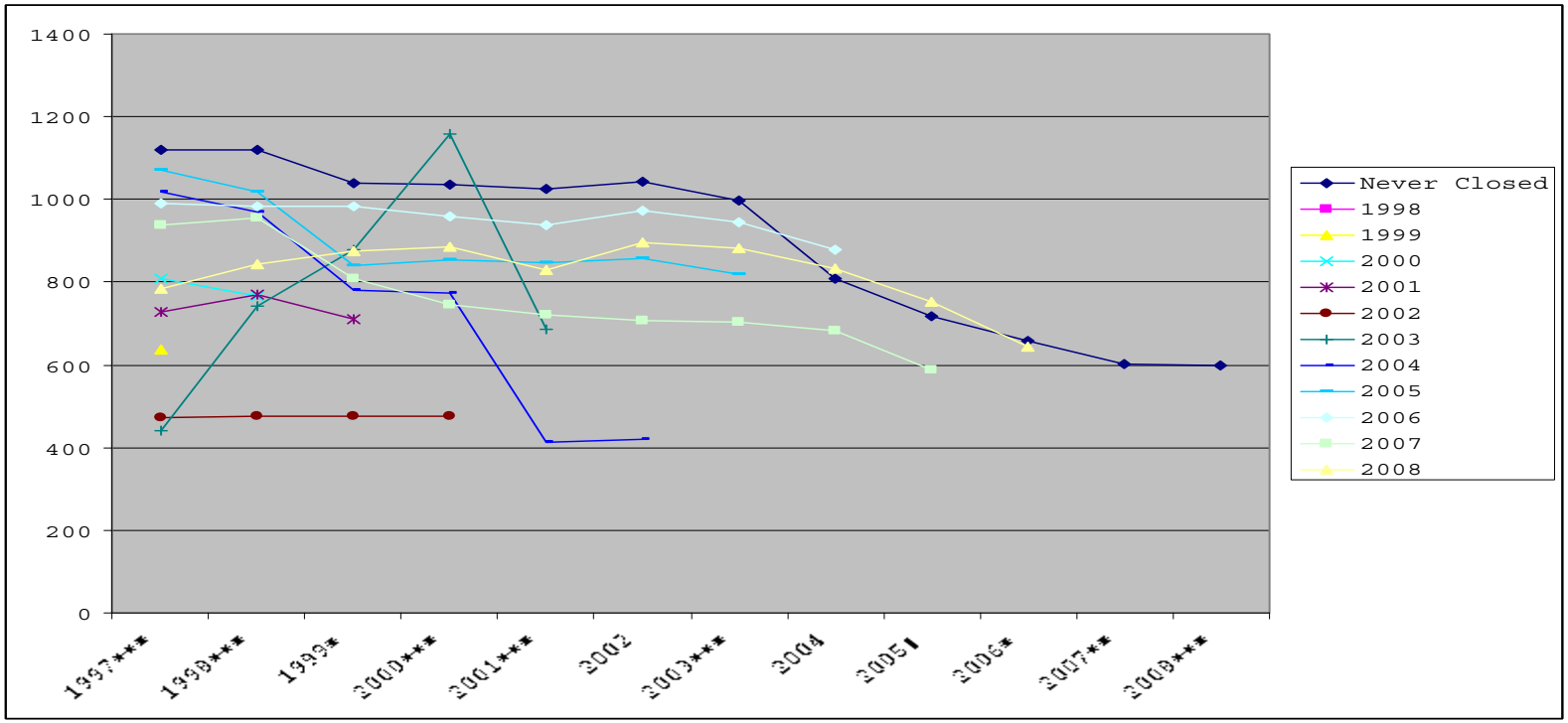


Figure 3-8. Enrollment by year of closure.

Table 3-6
Mean Growth in Enrollment from Prior Year, by Year of Closure

Last year of operation	1998***	1999***	2000***	2001***	2002*	2003**	2004***	2005***	2006***	2007***	2008*
Never Closed	10	24	41	0	37	-15	-40	-33	-26	-14	-8
1998	-20	Closed									
1999	-72	-161	Closed								
2000	-40	-234	-241	Closed							
2001	42	-58	-15	-179	Closed						
2002	3	2	-1	-144	-177	Closed					
2003	304	135	280	-474	5	-342	Closed				
2004	-49	-190	-8	-291	37	-56	-217	Closed			
2005	-54	-31	17	-8	10	-40	-278	-286	Closed		
2006	-5	-2	-24	-21	35	-27	-69	-314	-315	Closed	
2007	16	52	64	-24	-13	-4	-19	-94	-227	-161	Closed
2008	59	30	11	-47	65	-13	-48	-83	-108	-250	-210

†p<.10; *p<.05; **p<.01; ***p<.001

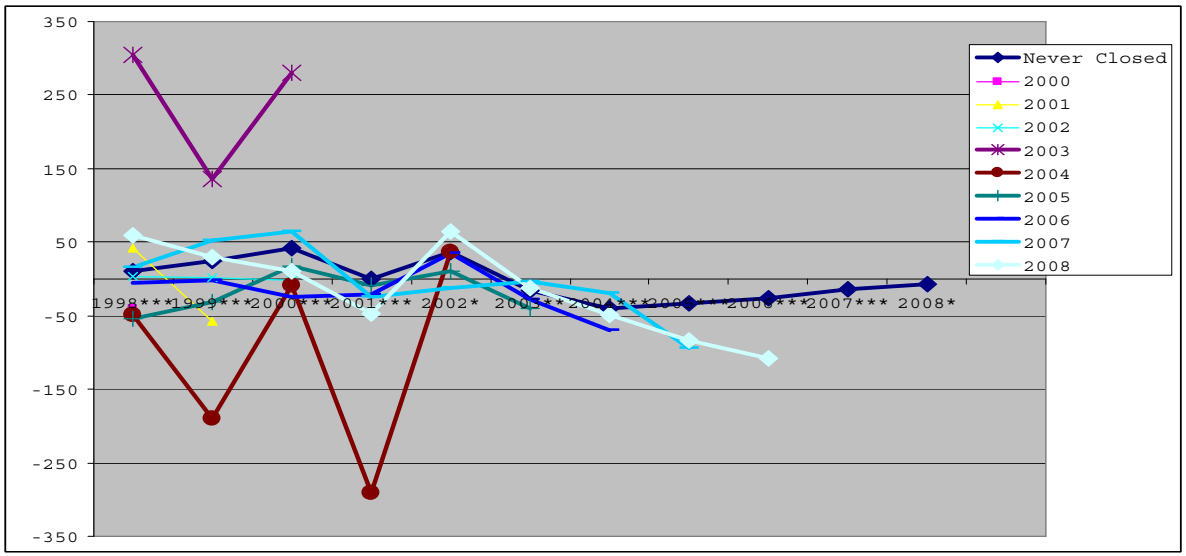


Figure 3-9. Growth in enrollment from the year prior, by year of closure.

Academic Characteristics of Students Who Attend Closed Schools prior to Enrollment

My findings thus far indicate that schools slated to be closed struggled academically for several years prior to closure. They also tended to experience small declines in enrollment. It is not clear, however, whether the association between closed schools and low average student performance was caused by the schools, or merely a reflection of the types of students that the to-be-closed schools initially enroll. A central criticism of closure policies is that schools identified for closure serve students who enter with greater needs and lower academic achievement than students in other schools. In this section I examine whether closed schools enrolled students who were more disadvantaged from the start, compared to other schools.

State test scores. Table 3-7 displays the school-average entering ELA scores for students entering the school by year of closure (e.g. fifth-grade scores for entering sixth graders for 6-8 middle schools; sixth-grade scores for entering seventh graders in 7-8 or 7-9 junior high schools). Not included in this table are entering scores for schools that are in the phase-out process (though very few enroll new students). Similar to school-average ELA scores while enrolled in the closed schools, from the start, these schools enrolled students who score significantly lower on average than students entering other middle schools. This disparity was true for both ELA and mathematics (see Table 3-8). The differences between school-average entering scores were similar to differences in school-average scores while enrolled. Average ELA scores for students entering to-be-closed schools were between 0.2 to 0.6 standard deviations below those of other schools. For the few years prior to closure, the gap was somewhat larger on average—between 0.3 to 0.6 standard deviations. The same trends were apparent in mathematics, although the largest gaps were closer to 0.5 standard deviations rather

than the 0.6 SD found with ELA scores. Across the entire decade of this study, these closed schools appear to have enrolled students who are at a significant disadvantage from the start, compared to other middle schools, particularly in the years prior to phase-out.

Table 3-7

Mean School-Average ELA Scores for Entering Students[a,b]

Last year of operation	School-average first-year students' entering (prior year) ELA scores								
	1997 scores (for 1998 entrants) ***	1998 (for 1999 entrants)*	1999 (for 2000 entrants)**	2000 (for 2001 entrants)**	2001 (for 2002 entrants)**	2002 (for 2003 entrants)**	2003 (for 2004 entrants)**	2004 (for 2005 entrants)†	2005 (for 2006 entrants)*
Never Closed	.167	.115	.148	.138	.113	.063	.018	-.035	-.062
1998	-.824								
1999									
2000	-.375	-.326							
2001	-.334	-.310	-.415						
2002	-.237	-.308	-.356	-.027					
2003	-.327	-.213	-.232	-.295					
2004	-.524	-.121	-.068	-.302	-.307	-.204			
2005	-.103	-.205	-.132	-.300	-.167	-.398			
2006	-.146	-.204	-.232	-.282	-.279	-.375	-.416	.188	
2007	-.338	-.204	-.235	-.272	-.321	-.467	-.394	-.465	-.511
2008	-.073	-.098	-.143	-.251	-.361	-.300	-.433	-.405	-.427

†p<.10; *p<.05; **p<.01; ***p<.001

[a]Blank cells indicate either the schools are phasing out or are no longer in operation.

[b] Measure is z-scored (M=0; SD=1) at the student level.

Table 3-8

Mean School-Average Mathematics Scores for Entering Students[a,b]

Last year of operation	School-average first-year students' entering (prior year) mathematics scores							
	1998 scores (for 1999 entrants)**	1999 (for 2000 entrants)**	2000 (for 2001 entrants) †	2001 (for 2002 entrants)**	2002 (for 2003 entrants)*	2003 (for 2004 entrants)**	2004 (for 2005 entrants)*	2005 (for 2006 entrants)†
Never Closed	.123	.151	.106	.116	.078	.026	-.012	-.052
1998								
1999								
2000	-.485							
2001	-.326	-.370						
2002	-.336	-.359	.094					
2003	-.267	-.267	-.171					
2004	-.131	-.112	-.380	-.304	-.124			
2005	-.204	-.174	-.277	-.159	-.317			
2006	-.209	-.218	-.301	-.319	-.292	-.419	.325	
2007	-.246	-.244	-.251	-.332	-.394	-.309	-.374	-.460
2008	-.132	-.183	-.237	-.318	-.300	-.377	-.324	-.519

†p<.10; *p<.05; **p<.01; ***p<.001

[a]Blank cells indicate either the schools are phasing out or are no longer in operation.

[b] Measure is z-scored (*M*=0; *SD*=1) at the student level.

Attendance rates. Similar disparities were evident in other academic characteristics. Tables 3-9 and 3-10 display mean school-average entering students' attendance rates by year of closure. With the exception of 2004 absence data for schools closed in 2006, average entering students' prior year absences were between one to five days greater in closed schools compared to other middle schools. Moreover, the gap in mean school-average absences between closed schools and other middle schools tended to increase in the few years prior to phase-out, suggesting that just prior to identification of a school for closure, schools tended to enroll more students with weaker prior attendance records. One exception to this trend are the schools closed in 2006, for which the gap in absences reversed closer to phase-out. Finally, schools closed in later years—2007 and 2008—had higher rates of absences and wider gaps between closed and other schools, compared to schools closed in earlier years.

Schools selected for closure also enrolled students who were more likely to be late to school in the years prior to entering the middle school (see Table 3-10). Mean school-average tardy rates for first year students in the year prior to enrollment were higher at closed schools by one to three days, compared to other middle schools. Additionally, in contrast to absences rates, most closed schools enrolled students with slightly lower rates of prior-year tardies in the year before phase-out. Differences across cohorts of closed schools were minimal.

Table 3-9

Mean School-Average Entering Students' Prior Year Absences[a]

Last year of operation	**School-average first-year students' prior year absences (days)**								
	2001 (for 2002 entrants)**	2002 (for 2003 entrants)		2003 (for 2004 entrants)		2004 (for 2005 entrants)		2005 (for 2006 entrants)	
	Spring	Fall	Spring**	Fall†	Spring**	Fall	Spring	Fall**	Spring***
Never Closed	6.24	4.66	6.22	4.70	5.80	4.70	6.69	4.85	6.43
2004	8.64	6.61	8.52						
2005	7.45	5.70	7.32						
2006	8.03	6.11	7.42	7.01	8.85	3.61	5.63		
2007	8.50	6.75	9.35	6.49	8.06	6.31	9.50	8.64	11.38
2008	8.04	5.57	7.38	6.22	7.85	5.59	8.23	6.58	8.93

†p<.10; *p<.05; **p<.01; ***p<.001

[a]Data are not available for attendance prior to spring, 2001. Blank cells indicate either the schools are phasing out or are no longer in operation.

Table 3-10

Mean School-Average Entering Students' Prior Year Tardies[a]

Last year of operation	**School-average first-year students' prior year tardies (days)**								
	2001 (for 2002 entrants)***	2002 (for 2003 entrants)		2003 (for 2004 entrants)		2004 (for 2005 entrants)		2005 (for 2006 entrants)	
	Spring	Fall***	Spring***	Fall***	Spring***	Fall*	Spring**	Fall	Spring***
Never Closed	3.43	4.24	4.30	3.98	4.41	3.55	4.30	4.09	4.75
2004	6.11	6.33	6.90						
2005	5.38	6.18	6.31						
2006	4.97	6.58	6.68	5.36	6.81	2.18	2.78		
2007	5.44	7.68	8.15	5.19	5.81	5.26	7.41	10.17	17.79
2008	5.50	5.03	5.83	5.16	5.46	4.87	5.62	5.25	6.25

†p<.10; *p<.05; **p<.01; ***p<.001

[a]Data are not available for attendance prior to spring, 2001. Blank cells indicate either the schools are phasing out or are no longer in operation.

Reconstituted Schools

New York City's approach to school closure usually involves creating and growing new schools in a building whose closing schools are phasing-out. I define these reconstituted schools those schools that either open newly in or move to the building housing the to-be-closed school when the closure is announced or in any future years following closure, and that serve sixth, seventh and/or eighth grade students. In several cases the closed schools had been sharing a school building with one or more schools for several years prior to the closure announcement. I did not categorize these pre-existing schools as reconstituting the closed schools. Additionally, I did not categorize as reconstituted any schools that moved to the building for only one year--only those that remained in the building for two or more years.

Within the same building, 134 schools replaced the 66 closed schools (see Table 6). The grade ranges reported here are the range of the school when first fully phased-in to the building—across the years, some of these schools expanded or contracted their grade spans.

Just over half (58%) of the newly reconstituted schools were middle or junior high schools (see Table 3-11). Another one-fifth of the reconstituted schools served middle-school-age students, with grade configurations of K-8, 6-12 and in one case, K-12. The other 20 percent of reconstituted schools included high and elementary schools, and one school that served only special education students.

Table 3-11
Grade Spans of Reconstituted Schools

Low grade	High grade	Number of schools
PK-5	3-6	10
5-8	7-9	78
PK-4	7-9	8
PK-4	10-12	1
5-8	10-12	18
9-10	11-12	16
Ungraded (special ed)		1
TOTAL		132

As displayed in Table 3-12, the majority of the schools that replaced the closed schools were traditional public schools. Charter schools did not begin moving into the buildings of closed middle schools until 2005; just under 10 percent of schools that replaced the closed schools were charter schools. Approximately twelve percent of the schools that moved into the buildings of the closed schools were not new schools, but rather had been incubated for a few years in another building, or in a few cases, were established over ten years or more prior to moving to the building. Additionally, many of the closed schools were already sharing their building space with other schools, and these remaining pre-established schools are not included in this count of reconstituted schools.

Table 3-12
Type of Reconstituted School in Buildings of Closed Schools

Year entered building	Charter	Traditional public	Other non-school (e.g. offices)
1995		2	
1996			
1997		4	
1998			1
1999		13	
2000		8	
2001		11	
2002		9	
2003		4	
2004		15	
2005	2	13	
2006	2	12	
2007	1	15	
2008	3	9	
2009	1	5	
2010	3		
TOTAL	12	120	1

One contention against school closure and reconstitution is that the new schools do not serve the same students that the closed schools served, and specifically, that the new schools select higher-performing students with lower academic needs. In the following section, I examine whether the reconstituted middle schools serve different types of students than the schools that were closed.

Reconstituted schools: Academic characteristics of entering students. Table 3-13 displays the mean school-average ELA and math scores for entering students in schools that were closed, compared to the new schools that reconstituted or replaced the closed schools. The table reports school-average scores for entering students two years prior to closure, which is approximately just prior to the closure announcement. For the reconstituted schools, the school-average entering scores were for the schools' first year of operation, usually one to two years after the closure announcement, and in two cases, three and six years later.

Similar to the schools selected for closure, school-average entering academic achievement characteristics at reconstituted schools were substantially lower than average. State test score differences between reconstituted and closed schools they replaced varied by year. For schools closed in 2000 through 2002—the years prior to the Bloomberg/Klein administration—reconstituted schools appear to have served equally or very slightly lower-achieving students. In later years, for schools closed in 2003 on, the reconstituted schools served slightly more advantaged, higher-achieving students compared to students who enrolled in the closed schools: for these years, school-average entering ELA and mathematics scores in the reconstituted schools were between .12 to .4 SD *higher* than their closed school counterparts, on average.

Table 3-13
Reconstituted Schools versus Schools Selected for Closure: School-Average Entering Student Test Scores

Year of closure	Entering (prior year) ELA scores [a, b]		Entering (prior year) mathematics scores [a, b]	
	Closed	Reconstituted	Closed	Reconstituted
2000 (ns = 9c, 10r)	-.375	-.446	n/a	n/a
2001 (n = 7c, 10r)	-.304	-.326	-.314	-.260
2002 (n = 3c, 3r)	-.356	-.320	-.359	-.352
2003 (n= 3c, 4r)	-.395	-.170	-.384	-.265
2004 (n = 4c, 5r)	-.307	.017	-.304	-.095
2005 (n = 10c, 14r)	-.398	-.221	-.317	-.173
2006 (n = 6c, 13r)	-.416	-.291	-.419	-.302
2007 (n = 5c, 7r)	-.465*	-.074	-.374†	-.063
2008 (n = 8c, 13r)	-.426**	-.242	-.519	-.302

†p<.10; *p<.05; **p<.01; ***p<.001
[a] Measure is z-scored (*M*=0; *SD*=1) at the student level.
[b] For closed schools, school-average scores are for entering students two years prior to closure. For reconstituted schools, the school-average entering scores are for the entering students in schools' first year of operation.

I find similar patterns with student academic behavior. Data are only available for schools closed in 2004 and later. Similar to patterns for entering academic achievement for schools closed after 2003, reconstituted schools enrolled student with fewer academic behavior problems compared to the closed schools they replaced (see Table 3-14). Specifically, reconstituted schools enrolled students who enter with slightly lower school-average absences in the prior year compared to schools selected for closure, on average. The differences in school-average entering absence rates were consistently lower in reconstituted schools, from .97 to 2.44 days depending on the year of closure. The pattern was similar for school-average tardiness of entering students: school-average prior year tardiness rates of entering students were between .3 to 1.57 incidences lower at the reconstituted schools, compared to the closed schools that they replaced.

Table 3-14
Reconstituted Schools versus Schools Selected for Closure: Entering Student Attendance Rates[a, b]

Year of closure	Prior spring school-average absences for entering students (mean)		Prior spring school-average tardies for entering students(mean)	
	Selected for closure	Reconstituted	Selected for closure	Reconstituted
2004 (n = 4c, 5r)	8.64†	6.31	6.11	4.75
2005 (n = 10c, 14r)	7.32†	6.35	6.31	4.99
2006 (n = 6c, 13r)	8.85	7.84	6.81	5.98
2007 (n = 5c, 7r)	9.50†	7.06	7.41	5.84
2008 (n = 8c, 13r)	8.93	7.58	6.25	5.95

†p<.10; *p<.05; **p<.01; ***p<.001
[a] Data are not available for attendance prior to spring, 2001.
[b] For closed schools, school-average attendance rates are for entering students two years prior to closure. For reconstituted schools, the school-average entering attendance rates are for the entering students in schools' first year of operation.

Reconstituted schools: Student demographics. Table 3-15 displays the demographic characteristics of reconstituted and closed schools by school year. Similar to differences in school-average achievement between reconstituted and closed schools, reconstituted schools

tended to serve a slightly lower percentages of students who are old for the grade, on average. Differences between reconstituted and closed schools ranged between three to eleven percentage points. There were two exceptions to this trend: for schools closed in 2001 through 2003, reconstituted schools served similar percentages of students who were old for their grade. With the exception of schools closed in 2006, reconstituted schools also served similar or very slightly lower percentages of mobile students on average compared to schools selected for closure, as measured by the number of moves a student made in elementary school.

No clear pattern emerged regarding student race and ethnicity, and across most years of closure, there appears to be very little displacement of students in terms of racial and ethnic background, with two exceptions. Reconstituted schools that replaced schools closed in 2003 served substantially higher percentages of black students and smaller percentages of white students than their closed counterparts, on average. For schools closed in 2003, average school racial composition was approximately one-quarter Black and one-quarter white. In contrast, the reconstituted schools were comprised of almost no white students (2.7%) and more than two-fifths (44%) black students. This change appears to have been driven by the closure of one school, Brooklyn Community School. Second, reconstituted schools that replaced schools closed in 2005 served slightly fewer black students and slightly more Hispanic students. Similarly, I find no clear relationship between reconstitution and home language, suggesting that reconstituted schools and the closed schools they replaced served similar percentages of English language learner students.

Table 3-15

Reconstituted Schools versus Schools Ever Selected for Closure: Demographic Characteristics by School Year[a]

Year of Closure	2000		2001		2002		2003	
	Ever closed (n=9)	Reconstituted (n=10)	Ever closed (n=7)	Reconstituted (n=10)	Ever closed (n=3)	Reconstituted (n=4)	Ever closed (n=3)	Reconstituted (n=4)
Racial composition (mean)								
Percent American Indian	0.3	0.2	0.2	0.5	0.2	0.8	0.4*	0.9
Percent Asian	1.3	1.2	3.0	1.1	2.6	1.4	8.3	1.7
Percent Black	40.6	42.9	34.9	35.8	51.6	51.1	25.8	44.1
Percent Hispanic	57.1	55.4	53.3	61.1	44.0	45.8	44.0	50.5
Percent White	0.8*	0.4	8.6	1.4	1.6	0.9	21.4	2.7
Percent eligible for free/reduced lunch (mean)	n/a	n/a	89.3	70.7	52.0	n/a	88.6	n/a
Percent old for grade (mean)	31.2*	20.1	16.4	17.7	22.0	18.5	13.1	16.1
Percent non-English household (mean)	8.7†	15.2	11.5	19.9	7.1	9.2	8.3	6.7
Percent with three or more moves in elementary school (mean)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

†p<.10; *p<.05; **p<.01; ***p<.001

[a] For closed schools, school-average demographics are two years prior to closure. For reconstituted schools, the demographics are for the schools' first year of operation.

Year of Closure	2004		2005		2006		2007		2008	
	Closed (n=4)	Reconstituted (n=5)	Closed (n=10)	Reconstituted (n=15)	Closed (n=6)	Reconstituted (n=13)	Closed (n=5)	Reconstituted (n=7)	Closed (n=8)	Reconstituted (n=13)
Racial composition (mean)										
Percent American Indian	0.7	0.7	0.5	0.4	0.6	0.7	0.9	0.4	0.9	0.5
Percent Asian	1.6	2.9	4.2	4.1	5.9	4.7	1.7	3.6	1.6	1.4
Percent Black	54.1	52.4	48.1	35.7	53.2	60.6	61.4	53.7	62.4	64.9
Percent Hispanic	41.6	40.0	45.5	55.5	37.1	31.3	33.4	38.2	32.1	29.7
Percent White	2.0	4.0	1.6	4.4	3.1	2.7	2.6	4.1	2.9	3.5
Percent eligible for free/reduced lunch (mean)	n/a	22.3	n/a	43.4	n/a	84.1	89.0	51.9	83.3	78.3
Percent old for grade (mean)	18.5	12.4	19.9	17.0	21.7	17.5	22.7	17.4	20.4	16.8
Percent non-English household (mean)	7.3	8.6	18.0	19.9	13.7	8.8	11.3	14.7	11.4	9.8
Percent with 3 or more moves in elementary school (mean)	30.6	30.5	28.7	26.8	30.9	32.5	35.8	28.5	29.6	24.4

†p<.10; *p<.05; **p<.01; ***p<.001

Closed and Reconstituted Schools Discussion

This description of closure and reconstitution in New York City from 1998-2008 provides some clarity regarding the characteristics of schools selected for closure and the actual implementation of closure and reconstitution. The findings here also provide some initial insights regarding the impact of closure on students. Proponents of school closure contend that the schools selected for closure are chronically underperforming and impenetrable to improvement. Critics of schools closure protest that the selection of schools is biased or politically motivated rather than based on a more objective rationale. My results suggest that descriptively the schools selected for closure did perform significantly worse than other middle schools on state assessments for multiple years prior to closure. Average absence and tardy rates were also higher in schools selected for closure compared to other middle schools. Additionally, these schools experienced small declines in enrollment over several years prior to closure, perhaps one market indicator of declining quality. Thus, it seems clear that schools selected for closure had significantly lower academic achievement, although there is enough variation in the school-average test scores to suggest the city does not base decisions solely on test scores. As such, the implementation of the closure policy may leave room for some political maneuverings, as some critics of closure charge. Relatively high-performing schools were closed in only a few cases, indicative that schools in New York City were closed mainly for accountability purposes and not solely for fiscal considerations as in many other school districts.

A second criticism of closure posits that the policy is racist in its targeting of schools in neighborhoods with high minority concentrations. The middle schools selected for closure did indeed serve a significantly higher percentage of Hispanic or Black students, compared to other middle schools: in earlier years, higher percentages of Hispanic students than other schools, and

in later years, higher percentages of Black students than in other schools. Charges of racist policies must confront the strong associations between school demographics and student achievement. In short, closing low-performing schools will result in closure of high-minority schools. I discuss these troubling associations later in this dissertation.

Closed schools also, not surprisingly given the long-documented links between poverty and school achievement, served higher percentages of poor students, as measured by free and reduced lunch eligibility. These schools also served higher percentages of mobile students and old-for-grade students, compared to other district middle schools. These findings point to the argument that closed schools perform poorly partly because they are serving some of the most vulnerable students in the District. Examination of the entering characteristics of first year students in closed schools compared to other middle schools provide evidence to support this argument. The mean gap in initial ELA and mathematics ability between closed and other schools was similar in magnitude to the gap between subsequent school-average scores. Furthermore, average entering students' prior year absences were greater in closed schools, and this difference is greatest in the period when the schools are identified for closure. Thus, these closed schools faced enormous challenges from the start in the enrollment and concentration of students with great academic need compared to other district schools.

Critics of the reconstitution element of the closure process also contend that any advantage reconstituted schools evidence after their counterparts close are due to the fact that the newly reconstituted schools replacing the closed schools tend to enroll more advantaged students. My results suggest that this assertion is at least partially true. The academic differences between the students who enroll in reconstituted schools compared the closed schools that they replaced are striking, particularly in later years. In 2004 and later, reconstituted schools

enrolled higher performing students, as measured by state test scores. Additionally, reconstituted schools enrolled students with fewer absences and tardiness in the year prior to enrolling, compared to the closed schools they replaced, suggesting that reconstituted schools enrolled students with slightly greater motivation, more supportive families or other social supports. Reconstituted schools also tended to serve slightly lower percentages of students who were old for the grade. In contrast, reconstituted schools were composed of relatively similar percentages of students who are mobile prior to entering middle school and who were eligible for free/reduced lunch, providing some evidence that the district was not particularly successful in altering the concentrations of poor students within the schools. Similarly, with few exceptions, reconstituted schools appear to have served similar types of students that the closed schools served in terms of race/ethnicity and home language.

These findings lend some credence to the argument that closure displaces students because in some respects the new schools are not serving the students and families who would have attended the closed school, although the changes in the academic composition of the student body may be purposeful on the part of the district in its stated goal to interrupt attendance patterns to lessen the high concentrations of low-achieving students. From a policy perspective, the District is faced with two conflicting choices: it can ensure that the reconstituted schools serve the same types of students as the closed schools they replace, reflective of the communities in which they are located, or the district can disrupt attendance patterns and create less academically and socio-economically segregated schools at the expense of no longer serving all students in the school community. I return to this policy dilemma later in this dissertation.

Chapter 4: Phase-Out Process

As stated earlier in this dissertation, New York City's approach to closure is somewhat exceptional in that it does not close a "failing" school and immediately open another. Rather, schools are phased out and phased in over several years. For example, once a school is identified for closure, the school will cease admitting new students that next fall, but allow currently enrolled students to continue, closing once the last group admitted of students has graduated. This next chapter examines how the city has implemented closure from 1998-2008, the characteristics of students as schools phase-out, and student mobility during the process.

Although these students represent quite a small percentage of New York City students, I believe the process is important to examine for two reasons. First, these are some of the most vulnerable and underserved students and thus warrant closer study. Second, a better understanding of student mobility during phase-out might provide some insight into the potential disruptions associated with school closure. Such information is important to my second research question regarding the impact of closure.

Implementation

Implementation of the closure process appears to vary across administrations, and in some cases, across schools. Of the 66 schools that closed, 13 did not go through a phase-out process two or three years prior to closure (see Table 4-1). In a few other cases, a school phased-out over a single year. For example, three middle schools that closed in 2001 appeared to implement only one year of phasing-out, leaving seventh graders without a school for eighth grade. Another fifth-eighth grade school dropped its fifth graders (who would have moved to

sixth grade in the final year), and sixth graders (who would have moved to seventh grade in its final year), retaining only the seventh graders, who moved up to eighth grade.

Table 4-1
Type of Closure Process

Last year of operation	Phased-out	Partially Phased-out (some but not all grades)	Did not phase-out	Total number of middle schools closed
1999-00	2		2	**4**
2000-01	8		1	**9**
2001-02		4	4	**8**
2002-03	2	1		**3**
2003-04	1		2	**3**
2004-05	4		1	**5**
2005-06	9		1	**10**
2006-07	5		1	**6**
2007-08	4		1	**5**
2008-09	8		0	**8**
Total				**64**

School Characteristics During Phase-Out

Table 4-2 displays the characteristics of schools as they went through the phase out process. Not unexpectedly, as schools phase-out grades, enrollment declined significantly. However, somewhat surprisingly, given that we would expect a negative impact of phase-out on achievement, the mean gap in school-average test scores between closing schools and all other schools appears relatively consistent, at approximately one-half standard deviation below the mean of other middle schools, across most years. School absence and tardy rates were also similar across most years, although schools experienced a very small rise in both absences and tardies in the fourth through second years prior to closure. School demographic characteristics appear to change as well: as schools approached closure, the percentage of students eligible for free and reduced lunch increased. Additionally, schools that were phasing out served primarily Black and Hispanic students—my findings suggest that very few White or Asian students experienced this process.

Table 4-2
School Phase-Out Process

	Five years prior to closure	Four years prior to closure	Three years prior to closure	Two years prior to closure	Last year of operation
Enrollment (mean)	799	787	689	502	284
School-average ELA scores [a] difference from mean of all other middle schools (mean)	-.485	-.523	-.558	-.558	-.530
School-average math scores[a] difference mean of from all other middle schools (mean)	-.486	-.503	-.550	-.544	-.536
School-average days absent (mean)					
Fall	8.3	8.7	8.7	9.0	8.4
Spring	10.7	11.5	11.4	11.1	10.8
School-average days tardy (mean)					
Fall	8.5	9.1	10.4	10.8	10.6
Spring	11.6	13.5	13.1	13.2	11.1
Racial composition (mean)					
Percent American Indian	0.5	0.5	0.5	0.5	0.51
Percent Asian	3.2	3.0	3.1	3.2	3.71
Percent Black	49.3	48.3	47.5	46.9	46.20
Percent Hispanic	41.1	44.1	44.9	45.5	45.73
Percent White	5.8	4.2	4.0	3.9	3.87
Percent eligible for free/reduced lunch (mean)	59.3	66.7	68.2	69.8	82.54
Percent non-English household (mean)	9.3	9.8	11.4	12.3	14.0

[a] Measure is z-scored (M=0; SD=1) at the student level.

Student Characteristics During Phase-Out

In my next set of analyses, I describe the characteristics of the students who attend the schools as they are phasing out, and focus on questions related to student mobility, including the extent to which students withdraw from the phasing-out school, how the characteristics of students who remain differ from to those who leave and enroll in other schools, and how the characteristics of these alternate schools to which students depart differ from the phasing-out schools. For these analyses, I exclude students enrolled in the 13 schools that did not phase-out prior to closure. Table 4-3 displays the percentage of students in grade five through eight that

experience closure and phase-out.[5] In any year, a very small percentage of students—between two and three percent of all middle school-age students—were attending middle schools that are in the process of closing.

[5] I am not able to include any data on phase-out and mobility between 8th and 9th grade for the junior high school that phased out and served 9th graders because I do not have 9th grade data for students in earlier years.

Table 4-3
Middle School-Age Students Experiencing School Phase-Out

	1998 (n=288,541)	1999 (n=289,474)	2000 (n=289,328)	2001 (n=298,054)	2002 (n=293,464)	2003 (n=284,315)
Other school (%)	97.4	96.5	97.4	98.2	99.4	97.1
Estimated year of announcement (%)	1.9	1.8	.4	.4	.1	2.7
First year of phase-out (%)	.7	1.2	1.5	.2	.4	.0
Second year of phase-out (%)	n/a	.5	.7	1.2	.1	.2

	2004 (n=276,716)	2005 (n=268,183)	2006 (n=264,983)	2007 (n=260,883)
Other school (%)	97.0	97.5	97.5	98.8
Estimated year of announcement (%)	1.4	1.0	1.5	n/a
First year of phase-out (%)	1.6	.8	.6	.9
Second year of phase-out (%)	.0	.7	.4	.3

Table 4-4 displays student mobility in schools that were not closed compared to schools that were closed, for middle-grade (5-8) students. Specifically, I examine the percentage of students who were enrolled in a school during the estimated year of announcement of closure and subsequently left the school during the phase-out process. I exclude from these mobility rates those students who left due to having reached the end of the grade configuration of the school. Although mobility rates differ significantly by phase-out status ($p<.001$), the relationship between phase-out and mobility was not consistent across years. In some years—1999, 2001, and 2002—mobility was considerably higher in schools phasing out. For example, in 1999, more than one-quarter (27.3%) of students attending schools in the year of announcement moved to a different school at some point during the phase-out, compared to one-fifth (20.4%) of students attending schools that were not phasing-out or about to phase-out. In contrast, in other years—2003 and later—mobility rates were the same or lower in schools phasing-out.

Table 4-4
Student Mobility by Closure Status

School year	Percentage of students who moved schools in the subsequent year			
	Not closing	Estimated year of announcement	First year of phase-out	Second year of phase-out
1998	21.7%***	13.8%	10.2%	n/a
1999	20.4%***	13.3%	7.5%	7.5%
2000	19.1%***	14.8%	12.2%	5.7%
2001	18.8%***	31.9%	12.0%	45.7%
2002	18.5%***	15.1%	30.3%	28.1%
2003	20.5%***	11.7%	3.1%	.0%
2004	19.7%***	6.9%	10.3%	4.5%
2005	18.6%***	8.4%	6.2%	9.6%
2006	18.4%***	13.1%	7.7%	1.6%
2007	20.6%***	n/a	16.1%	1.9%

Years with particularly high mobility rates during the phase-out process—approximately one-third and larger—are usually because one school dropped a grade completely while closing.

For example, note the particularly high mobility rate (31.9%) for schools in the year of announcement in 2001 and subsequently, the first year of phase-out in 2002 (30.3%). One school phasing-out during these year enrolled sixth through ninth graders in the estimated year of announcement, sixth through eighth graders in the first year of phase-out, and only eighth graders in the year prior to closure. Thus, all eighth (and graduating ninth) graders moved to a new school following the year of announcement, and all sixth graders moved to a new school following the first year of the phase-out. Similarly, the schools in the second (and last) year of phase-out in 2001 enrolled both 7th and 8th graders, thus the following year all seventh graders were required to move to new schools.

Students leaving closing schools moved to schools with higher average achievement scores and fewer attendance problems, on average. Table 4-5 displays the characteristics of the phasing out schools compared to the schools that receive students from these schools the next year. Characteristics for both schools are in the estimated year of closure announcement. With the exception of 2001, across all years the schools that students moved to had higher school-average ELA and mathematics scores; receiving schools had school-average ELA and mathematics scores that were 0.2 to 0.5 SDs higher, on average, than those of the closing schools. The schools that students moved to also had significantly lower average tardy rates; school-average tardy rates for closed schools were between three to six days higher in closed schools across the years of this study. School-average absence rates in 2001 and 2002 were slightly higher at receiving schools, on average, but in 2003 and later, school-average absence rates at closed schools were between one-half to two days higher compared to receiving schools.

Table 4-5

School Academic Characteristics of Phasing Out and Receiving Schools

School year	**1998-99**		**1999-2000**		**2000-01**		**2001-02**		**2002-3**	
Characteristics	Phasing out (n=7)	Receiving Schools (n=105)	Phasing out (n=9)	Receiving Schools (n=110)	Phasing out (n=2)	Receiving Schools (n=195)	Phasing out (n=3)	Receiving Schools (n=91)[a]	Phasing out (n=1)[c]	Receiving Schools (n=52)
School-average ELA scores[a] (mean)	-.637**	-.245	-.442†	-.194	-.638*	-.185	-.298	-.335	-.400	-.255
School ELA standard deviation (mean)	.919	.903	.912	.894	.911	.896	.807	.875	.857	.802
School-average math scores[a] (mean)	-.639***	-.232	-.463*	-.199	-.629†	-.204	-.205	-.319	-.317	-.264
School math standard deviation (mean)	.930	.890	.959†	.882	.927	.886	.831	.862	.898	.854
School-average days absent (mean)										
Fall	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	8.7	9.22
Spring	n/a	n/a	n/a	n/a	n/a	n/a	10.1	13.1	10.7	11.1
School-average days tardy										
Fall	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	11.6	6.7
Spring	n/a	n/a	n/a	n/a	n/a	n/a	11.6	8.0	14.0	8.0

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Measure is z-scored (*M*=0; *SD*=1) at the student level.

[b] Student test scores, attendance, mobility, language and age data for JHS 142 Stranahan are not available.

[c] Data for JHS 17 LaSalle School are not available.

School year	2003-4		2004-5		2005-6		2006-7	
Characteristics	Phasing out (n=13)	Receiving schools (n=149)	Phasing out (n=5)	Receiving schools (n=169)	Phasing out (n=5)	Receiving schools (n=142)	Phasing out (n=5)	Receiving schools (n=149)
School-average ELA scores [a] (mean)	-.416*	-.184	-.603*	-.257	-.470	-.311	-.489†	-.166
School ELA standard deviation (mean)	.827	.861	.865	.833	.859	.852	.811	.837
School-average math scores [a] (mean)	-.343†	-.146	-.635*	-.239	-.561	-.316	-.527†	-.227
School math standard deviation (mean)	.884	.882	.951	.863	.872	.851	.787	.832
School-average days absent (mean)								
Fall	8.6	8.0	8.9	7.8	9.5	9.7	8.6	7.2
Spring	10.7	9.7	12.4	10.7	12.2	11.6	11.2	9.7
School-average days tardy								
Fall	10.4*	6.8	9.4*	5.9	9.1	6.4	13.6**	7.9
Spring	13.5***	8.5	12.8*	8.1	14.7*	8.6	15.3*	9.8

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Measure is z-scored (*M*=0; *SD*=1) at the student level.

With respect to demographic characteristics during the phase-out process, the schools that students from closing schools move to served a population of students that is less disadvantaged and less segregated. Receiving schools were larger than the closing schools, by between 100 to 800 students, on average (this difference in size is not due to phase-out; characteristics for both schools were in the estimated year of closure announcement; see Table 4-6). Additionally the racial/ethnic composition of the student body of the receiving schools was slightly less segregated with lower concentrations of any one ethnic group. The same was true for poverty. With the exception of the 2000 school year, receiving schools tended to have slightly lower proportions (5-10%) of students eligible for free or reduced lunch, on average. The receiving schools also had lower percentages of students who were old for their grade and lower percentages of students with high rates of mobility prior to entering middle school. This pattern is similar to the closure-reconstitution process in which the non-closure schools enrolled higher-achieving students, although the reconstituted schools were similarly racially and socio-economically segregated.

Table 4-6
School Demographic Characteristics of Phasing Out and Receiving Schools

School year Characteristics	**1998-99** Closed: last year is 2000-1 (n=7)	Receiving Schools (n=105)	**1999-2000** Closed: last year is 2001-2 (n=9)	Receiving Schools (n=110)	**2000-01** Closed: last year is 2002-3 (n=2)	Receiving Schools (n=195)	**2001-02** Closed: last year is 2003-4 (n=3)	Receiving Schools (n=91)[a]	**2002-3** Closed: last year is 2004-5 (n=1)[c]	Receiving Schools (n=52)
Enrollment (mean)	845	1055	658*	1068	641	1102	498	1352	292	1106
Racial composition (mean)										
Percent American Indian	0.3	0.3	0.3	0.3	0.2	0.3	0.3	0.3	1.0	0.4
Percent Asian	1.5	5.8	2.4	6.9	0.1	4.5	7.4	5.6	3.4	8.7
Percent Black	39.8	43.9	35.8	43.9	61.5	47.9	26.2	38.5	50.0	38.3
Percent Hispanic	57.7	43.4	54.6	41.0	37.4	42.1	47.8	50.7	40.1	43.5
Percent White	0.8***	6.7	6.8	8.0	6.9	5.2	18.2	4.9	5.5	9.2
Percent eligible for free/reduced lunch (mean)	n/a	n/a	94.8*	82.9	44.6	79.7	89.4	83.9	n/a	n/a
Percent old for grade (mean)	31.2**	19.5	18.6†	13.5	25.3*	14.4	17.4	17.7	13.8	17.6
Percent non-English household (mean)	8.7	6.1	10.8	7.8	6.9	9.4	9.3	11.4	6.7	14.0
Percent with three or more moves in elementary school (mean)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	28.6	21.5

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Student test scores, attendance, mobility, language and age data for JHS 142 Stranahan are not available.

[b] Data for JHS 17 LaSalle School are not available.

School year	2003-4		2004-5		2005-6		2006-7	
Characteristics	Closed: last year is 2005-6 (n=13)	All other MS (n=149)	Closed: last year is 2006-7 (n=5)	All other MS (n=169)	Closed: last year is 2007-8 (n=5)	All other MS (n=142)	Closed: last year is 2008-9 (n=5)	All other MS (n=149)
Enrollment (mean)	728†	1075	934	908	589	1133	655	777
Racial composition (mean)								
Percent American Indian	0.5	0.4	0.7	0.5	0.9	0.6	0.6	0.6
Percent Asian	3.6	7.4	0.8	5.0	1.7	6.2	0.9	5.8
Percent Black	49.4	40.3	58.8	43.2	61.4	49.2	64.6	47.8
Percent Hispanic	45.1	44.8	39.0	45.3	33.4	39.6	33.2	39.9
Percent White	1.5***	7.0	0.7	6.0	2.6	4.5	0.7	5.9
Percent eligible for free/reduced lunch (mean)	n/a	n/a	n/a	n/a	89.0	79.9	88.1	88.1
Percent old for grade (mean)	20.2**	15.2	24.7**	16.2	22.7**	15.1	20.5†	15.3
Percent non-English household (mean)	15.6	15.5	13.8	17.0	11.3	15.7	12.6	16.0
Percent with three or more moves in elementary school (mean)	30.4†	25.1	33.4†	26.4	35.9**	26.0	29.5	25.9

†p<.10; *p<.05; **p<.01; ***p<.001

As discussed earlier in this chapter, school-average academic characteristics during phase-out remained relatively constant, suggesting perhaps that the students who chose to leave schools during the phase-out process were not much different than students who elected to stay, and further, that the experience of phasing out did not impact students. My examination of student-level characteristics suggests, however, some differences across students who experienced phase-out and students who moved prior to phase-out (see Table 4-7).[6]

Although students in both groups had extremely low test scores, close to -0.5 SD on average, students who left during phase-out had slightly higher average test scores across all grades, compared to students who remained in the closing schools. The gap was approximately the same size across grades, and was apparent both prior to entering middle school as well as during middle school. In contrast, however, these students who left had higher rates of absences and tardies than students who stayed, and these differences in the number of absences and tardies grew wider in later grades. For example, the average fall semester absence rate for sixth-grade students who remained was about one day less than students who left. By eighth grade, the average fall semester absence rate for eighth-grade was about three days lower, on average. Furthermore, although not displayed in this table, the comparatively stronger academic characteristics of school movers was not reflective of trends in the district population as a whole; across all middle schools, students who moved schools rather than remain tended to have lower state test scores and higher rates of absences and tardies. Finally, regarding demographic characteristics, movers and stayers appear similar, with one exception: students who stayed enrolled in the phasing-out school were less likely to Black and more likely to be Hispanic.

[6] In the **remained** group I include students who were in the school at the time of announcement in grade five, six and seven and never moved, meaning if they were in grade seven they experienced one year of phase-out. If they were in grade six, two years. In the **moved** group, I include students who were in the school at the time of announcement in grade five, six and seven and ever moved – in the first or second year of phase-out. As such, some of these students (the fifth and sixth graders) may have experienced one year of phase out.

Table 4-7

Characteristics of Students Who Leave and Stay during Phase-Out

	Remained	Moved
ELA scores [a] (mean)		
Grade 4	-.431	-.459
Grade 5	-.418***	-.479
Grade 6	-.445***	-.526
Grade 7	-.475*	-.516
Grade 8	-.462*	-.496
Math scores[a](mean)		
Grade 4	-.381***	-.447
Grade 5	-.390***	-.479
Grade 6	-.455***	-.551
Grade 7	-.479***	-.577
Grade 8	-.492***	-.565
Days absent (mean)		
Grade 4 Fall	6.1***	7.0
Grade 4 Spring	7.3***	8.4
Grade 5 Fall	5.8***	6.8
Grade 5 Spring	7.7***	8.8
Grade 6 Fall	6.5***	8.2
Grade 6 Spring	8.7***	11.2
Grade 7 Fall	7.5***	10.7
Grade 7 Spring	9.8***	13.6
Grade 8 Fall	8.1***	10.9
Grade 8 Spring	10.9***	15.0
Days tardy (mean)		
Grade 4 Fall	5.2***	6.6
Grade 4 Spring	4.9***	6.3
Grade 5 Fall	5.1***	6.7
Grade 5 Spring	5.4***	6.9
Grade 6 Fall	6.5***	7.5
Grade 6 Spring	9.0**	9.8
Grade 7 Fall	9.8***	10.8
Grade 7 Spring	12.5	12.6
Grade 8 Fall	11.7***	10.7
Grade 8 Spring	12.7**	13.6
Race/ethnicity (%)***		
American Indian	.5%	.6%
Asian	3.3%	2.8%
Black	43.8%	50.4%
Hispanic	49.8%	43.9%
White	2.4%	1.8%
Multiracial, other	.2%	.5%
Eligible for free/reduced lunch(%)	34.0%	31.4%
Old for grade (%)	20.7%*	22.2%
Non-English household (%)	14.5%	14.9%
Number of elementary schools (mean)	1.93***	2.21
Special education (%)	10.8%	10.8%

†p<.10; *p<.05; **p<.01; ***p<.001

[a] Measure is z-scored (*M*=0; *SD*=1.

One argument against closure contends that during phase-out and after closure, nearby schools that may already be overburdened receive an influx of students due to the closure. Table 4-8 displays the number of students transferring from phasing-out schools into other receiving schools. Students who moved during phase-out moved to a number of schools; most years schools received a mean of less than five students overall, although from 2002 to 2004, the influx of students received was between 15 and 20 students on average. However, the number of students enrolled varied widely, as the large standard deviations suggest. As displayed in Figure 4-1 most schools received fewer than 25 students per year. However, during 2002 through 2005, a few schools received more than 50, and in a couple of cases, close to 200 students in one year.

Table 4-8
Receiving Schools

School year	Number of schools phasing out in the current school year	Number of schools receiving PO students	Minimum number of PO students received	Maximum number of PO students received	Mean number of PO students received	SD
1998	4	38	1	4	1.37	.7
1999	11	124	1	20	3.20	3.5
2000	16	120	1	32	2.30	3.2
2001	9	105	1	30	2.57	3.9
2002	5	97	1	175	7.47	20.5
2003	4	55	1	87	6.00	14.9
2004	14	169	1	158	4.03	15.2
2005	15	181	1	54	2.51	4.4
2006	10	152	1	34	2.30	4.1
2007	13	154	1	22	1.97	2.5
2008	8[a]	42	1	6	1.26	.9

[a] Data on phase-out status are available only for schools in the second year of phase-out; thus excluded from 2008 counts are schools in the first year of phase-out.

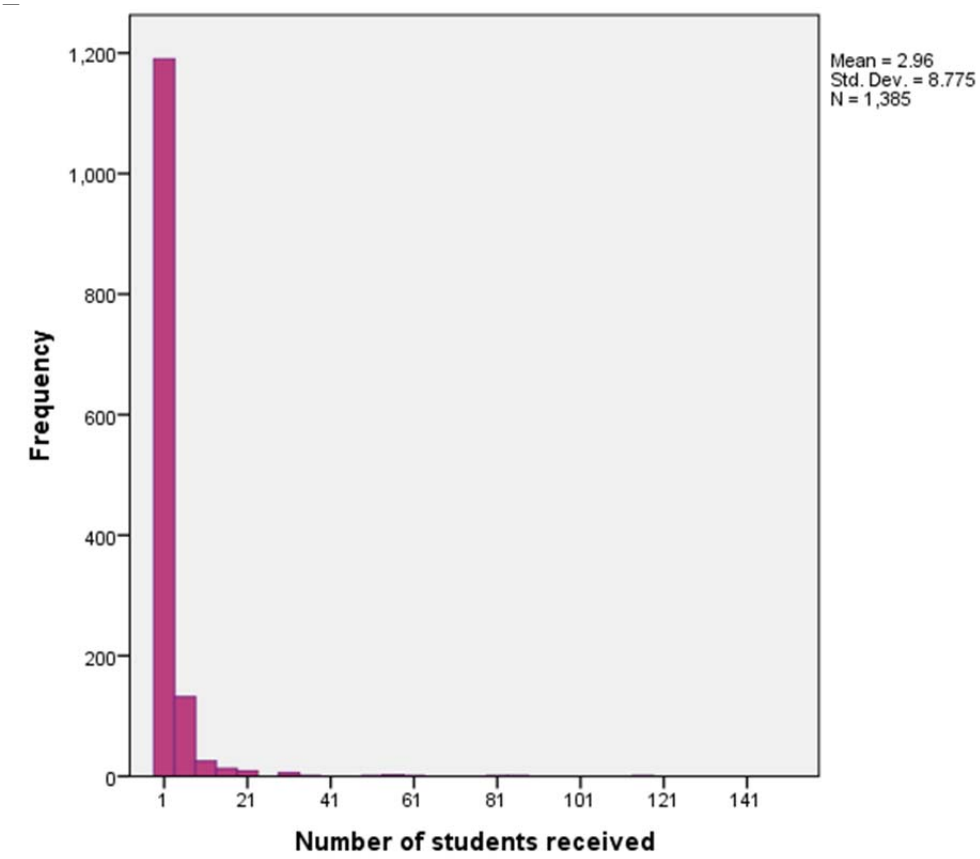


Figure 4-1. Number of students received from closing schools.

Phase-Out Process: Discussion

The gradual nature of New York City's closure process may allow for increased personal attention to students as the schools get smaller, which in turn may serve to minimize student mobility. On the other hand, the stigma associated with closure and the extended nature of the phase-out process might result in greater student mobility. Critics of closure also argue that schools lose important academic resources during phase-out as they grow smaller and prepare to close.

The variability in mobility rates across years during phase-out suggests that changing District policies may play a role in determining mobility rates during phase-out, particularly given that the differences align with changes in District administrations. Anecdotal reports suggest that nearby schools often receive an influx of higher-need students when a neighboring school is phasing out and closed. My analysis suggests that in a few cases in certain years, a few schools were burdened by the enrollment of 50-200 students moving from closing schools, but that in most years, schools enrolled very few students fleeing the phase-out process. Thus, phasing schools out rather than closing schools immediately, an approach often taken by other large districts such as Chicago, appears to lessen the impact on neighboring schools.

Students who move during phase-out had slightly higher test scores on average, compared to students who stayed. In contrast, movers had higher rates of absences and tardies than students who remained. The differential findings between the test scores trends and attendance trends are puzzling. The test score differences might partially be an artifact of the state tests, which are not necessarily optimal instruments to measure

achievement for students that are already performing quite low. For example, by definition mean reversion is more pronounced among low-performing students.

This chapter's findings also provide some direction for my second research question on the impact of closure. The trends are suggestive, particularly the fact that students who left school during phase-out ended up at higher-performing schools. But a closer analysis of the impact of closure is required. Analysis of impacts should examine effects by year, given the differential rates of mobility across administrations. Measures of the impacts of closure should also take into account the blunt nature of the state test scores, by considering potential ceiling or floor effects, and by using other measures of academic behavior such as attendance rates.

Chapter 5: Impact of School Closure

My analysis of the impact of closure focuses on middle schools (grades 6-8) whose last year of operation was in 2006, 2007 or 2008. The NYC DOE closed fourteen middle schools during this three-year period. The descriptive results in the prior chapter indicate that three years prior to closure, these failing schools served substantially larger proportions of academically and socially at-risk students. It is therefore possible that the poor academic performance of schools selected for closure simply reflected the types of students they served, rather than the contribution these schools made to their students' learning and attendance. My primary analytic models attempt to disentangle selection effects from causal effects, comparing academic growth among students who enrolled in sixth grade in schools that would soon be closed to future cohorts of students who would have enrolled in the closed schools had they not been closed.

Difference-in-Differences Analyses

Table 5-1 presents my findings regarding the impact of enrolling in a chronically failing school on student academic development in ELA and mathematics, based on my matching, difference-in-differences analyses. All estimates are fully adjusted for the available student-level covariates. As indicated by the intercept for each model, students in both groups entered middle school at a considerable academic disadvantage compared to students who did not attend, or were not predicted to have attended, a school slated for closure. Reflecting the school-level findings from the previous chapter, average students' fifth grade ELA and mathematics test scores were roughly one-third to one-half standard deviation below the New York City average during each time period.

As the grade by to-be-closed-school interaction term coefficients indicate, within each cohort and subject, students who enrolled in chronically underperforming schools learned slightly less over the course of their middle school careers compared to students who attended alternate receiving schools. These negative effects varied by grade and by cohort. School-closure estimates indicate negative effects in seventh grade associated with enrollment at a to-be-closed school, but not in other grades. There were two exceptions to these patterns: I found no differences in ELA test scores for students who enrolled in schools closed in 2007 compared to their counterparts who attended other schools, although negative effects in mathematics were evident for students who enrolled in schools closed this year. I also found no differences in math test scores for students who enrolled in schools closed in 2006, although in contrast, I found negative effects on ELA test scores for students who enrolled in schools closed in 2006.

Table 5-1
Impact of Enrolling in a School Identified for Closure on ELA and Mathematics Learning

	ELA	Mathematics
2006 (n=5,242)		
Intercept	-0.457*	-0.510*
	(0.023)	(0.023)
Enrolled in a to-be-closed school	-0.004	-0.003
	(0.025)	(0.026)
Sixth grade	0.016	0.017
	(0.022)	(0.022)
Sixth grade* to-be-closed school	-0.104*	-0.015
	(0.031)	(0.033)
Seventh grade	0.087*	0.037*
	(0.013)	(0.014)
Seventh grade* to-be-closed school	-0.056*	-0.029
	(0.018)	(0.020)
Eighth grade	-0.006	-0.028
	(0.015)	(0.017)
Eighth grade* to-be-closed school	-0.018	-0.038
	(0.019)	(0.022)
2007 (n=1,974)		
Intercept	-0.409*	-0.389*
	(0.046)	0.044
Enrolled in a to-be-closed school	0.005	0.003
	(0.042)	(0.041)
Sixth grade	0.082	0.051
	(0.054)	(0.047)
Sixth grade* to-be-closed school	-0.019	0.093
	(0.060)	(0.051)
Seventh grade	0.057*	0.072*
	(0.026)	(0.025)
Seventh grade* to-be-closed school	-0.043	-0.093*
	(0.037)	(0.031)
Eighth grade	0.068	0.007
	(0.037)	(0.032)
Eighth grade* to-be-closed school	-0.080	-0.058
	(0.046)	(0.036)

*p<.05; **p<.01; ***p<.001.

	ELA	Mathematics
2008(n=3,890)		
Intercept	-0.413*	-0.479*
	(0.028)	(0.028)
Enrolled in a to-be-closed school	-0.048	0.020
	(0.028)	(0.032)
Sixth grade	0.066	0.041
	(0.030)	(0.022)
Sixth grade* to-be-closed school	-0.003	0.042
	(0.033)	(0.039)
Seventh grade	0.076*	0.117*
	(0.016)	(0.017)
Seventh grade* to-be-closed school	-0.123*	-0.142*
	(0.023)	(0.023)
Eighth grade	0.025	-0.049
	(0.024)	(0.025)
Eighth grade* to-be-closed school	0.012	0.041
	(0.029)	(0.030)

*$p<.05$

With ELA, in schools closed in 2006 and 2008, children attending failing schools gained somewhat less in seventh grade (ES=-0.06 and -0.12, respectively), but had parallel growth rates in sixth and eighth grades compared to similar students who attended alternate schools. No effects were found for the 2007 cohort in ELA. Similarly, with mathematics, differences between failing and alternate schools also vary by grade and cohort. With both the 2007 and 2008 cohorts, students attending schools that would soon be closed gained slightly less in mathematics in seventh grade (ES=-0.09 and -0.14, respectively), while no effects were found for the 2006 cohort. Importantly, however, note that across all cohorts, subjects, and grades, students attending both the closed and alternate schools scored well below other New York City students. Figures 5-1 to 5-3 provide graphic displays of the estimated effects of closure by the time students reach eighth grade, organized by subject, based on the models displayed in Table 4-1. The y-

axis indicates achievement levels in standard deviation units compared to all students in New York City in a given year, grade, and subject.

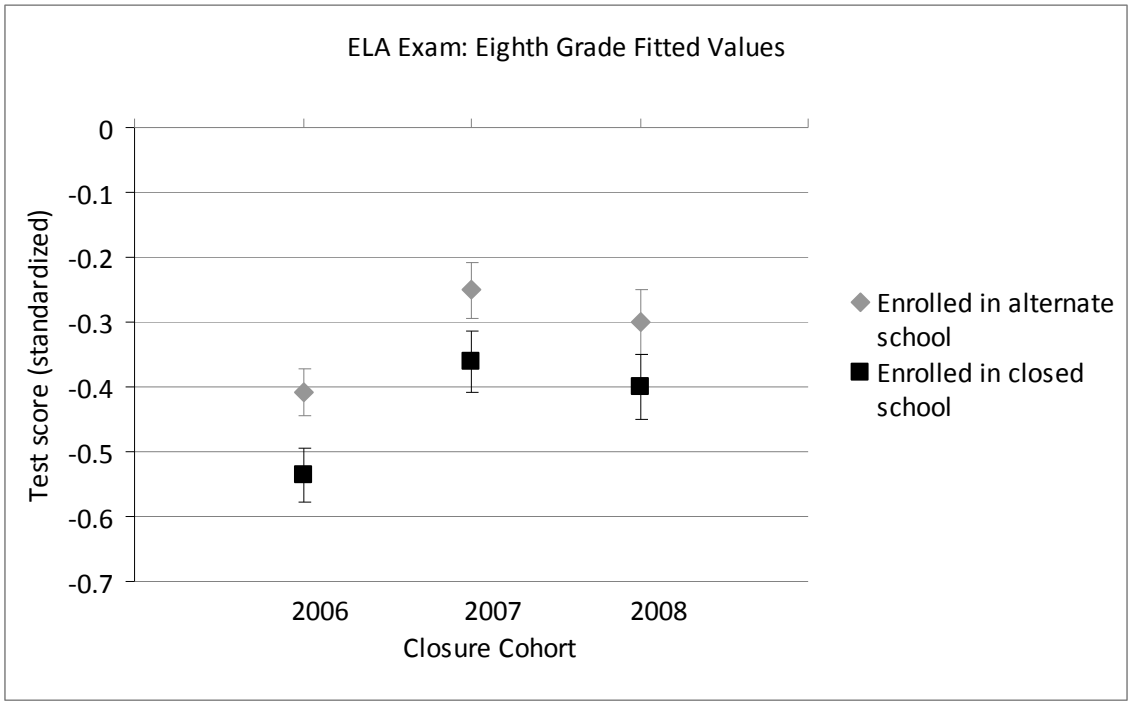


Figure 5-1. Closure: ELA exam.

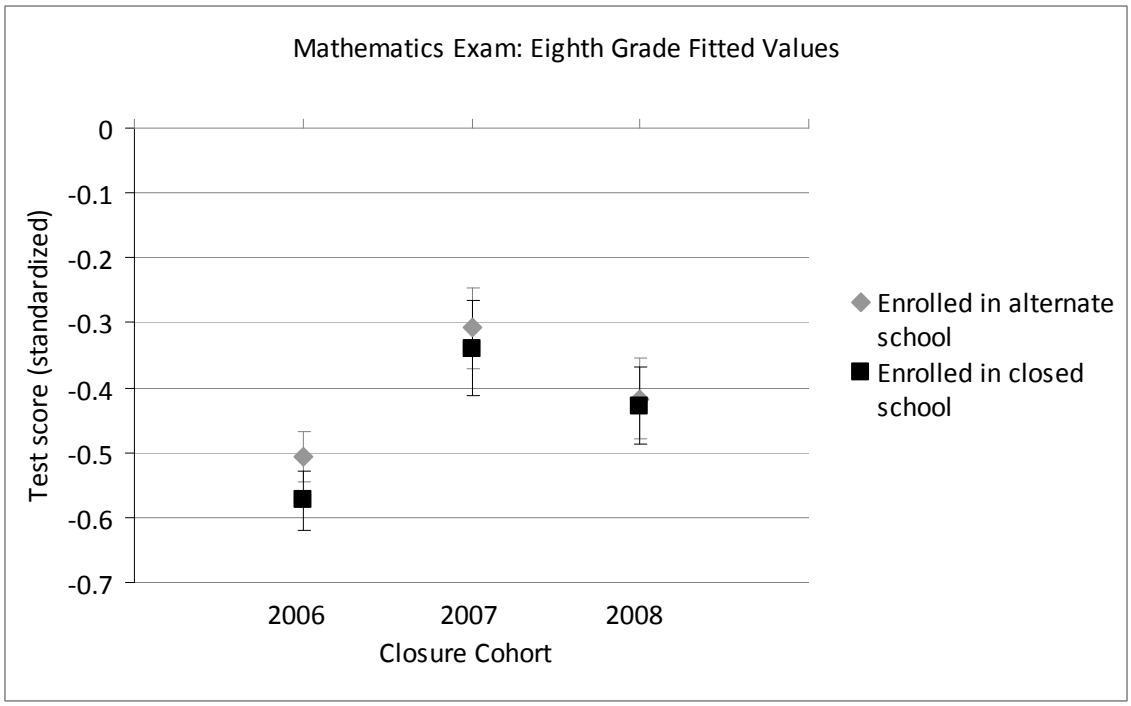


Figure 5-2. Closure: Mathematics exam.

Absences

In addition to examining student test score trajectories, I also examined the impact of school closure on student absences. I include absences as a second academic outcome in part as a robustness check on test scores, because absences are less likely to have measurement error and mean reversion. Second, absences themselves are an important indicator of student achievement—being absent from school is associated with a decrease in learning and achievement, lower exam scores, higher rates of dropping out, drug and alcohol use and unemployment (Gottfried, 2009).

Table 5-2 presents my findings regarding the impact of enrolling in a chronically failing school on student absences (log transformed). All estimates are fully adjusted for the available student-level covariates. Students in both groups have an average of between 2.4 to 2.5 logged absences (11 to just under 13 absences) in the year prior to entering middle school, as indicated by the intercept in each of the three cohort models.

The grade by to-be-closed-school interaction term coefficients indicate that within each cohort and subject, students who enrolled in chronically underperforming schools have higher rates of absences over the course of their middle school careers compared to students who attended alternate receiving schools. These negative effects varied by grade and by cohort. All students in both treatment and control have higher absences in eighth grade. Four of the nine school-closure estimates—three cohorts, three grades per cohort—indicate negative effects associated with enrollment at a to-be-closed school.

Table 5-2
Impact of Attending a School Identified for Closure on Logged Absences

	Absences (logged)
2006 (n=5,242)	
Intercept	2.546*
	(0.026)
Enrolled in a to-be-closed school	0.038
	(0.021)
Sixth grade	-0.001
	(0.022)
Sixth grade* to-be-closed school	0.011
	(0.043)
Seventh grade	0.032
	(0.019)
Seventh grade* to-be-closed school	0.067*
	(0.022)
Eighth grade	0.166*
	(0.015)
Eighth grade* to-be-closed school	-0.012
	(0.020)
2007 (n=1,974)	
Intercept	2.529*
	(0.046)
Enrolled in a to-be-closed school	-0.115*
	(0.045)
Sixth grade	-0.070
	(0.038)
Sixth grade* to-be-closed school	0.040
	(0.060)
Seventh grade	0.033
	(0.023)
Seventh grade* to-be-closed school	0.078*
	(0.033)
Eighth grade	0.118*
	(0.024)
Eighth grade* to-be-closed school	0.086*
	(0.034)

	Absences (logged)
2008 (n=3,890)	
Intercept	2.410*
	(0.036)
Enrolled in a to-be-closed school	-0.062
	(0.034)
Sixth grade	0.023
	(0.039)
Sixth grade* to-be-closed school	0.015
	(0.045)
Seventh grade	0.021
	(0.022)
Seventh grade* to-be-closed school	0.087*
	(0.028)
Eighth grade	0.097*
	(0.025)
Eighth grade* to-be-closed school	0.017
	(0.030)

*$p<.05$

Figure 5-3 provides graphic displays of the fitted estimates in Table 5-2, organized by year of closure. The y-axis indicates number of days absent per year in a given grade (in the non-logged, days metric). For schools closed in 2006 and 2008, children attending chronically underperforming schools had higher rates of absences in seventh grade (0.07 and 0.09 logged absences, respectively, a little over one day higher), but parallel rates in sixth and eighth grades compared to similar students who attended alternate schools. For schools closed in 2007, students attending schools that would later be closed had higher rates of absences in seventh and eighth grade (0.08 and 0.09 logged absences, respectively), and parallel rates of absences in sixth grade compared to their counterparts in alternate schools. It is interesting to note, across all academic outcomes, seventh grade is a common inflection point.

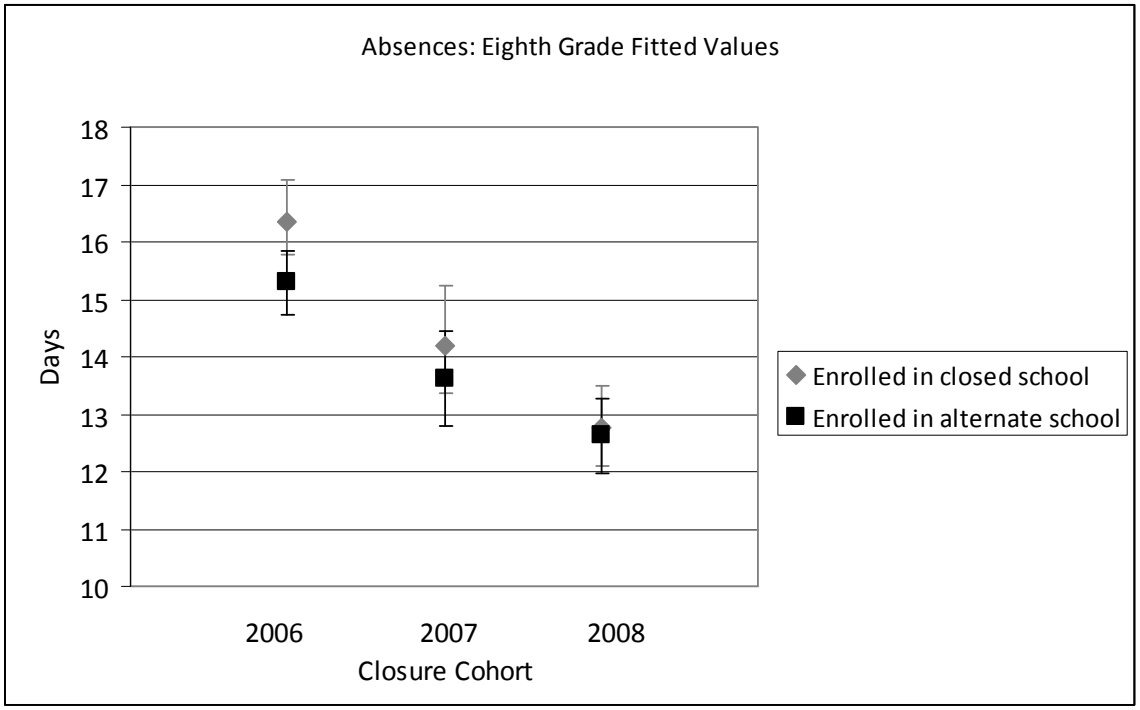


Figure 5-3. 2006 Closure: Absences.

Difference-in-Differences Discussion

The middle schools New York City closed during the mid-2000s for poor performance did indeed have very weak student test scores, and were isolated both economically and racially/ethnically. The findings from my primary analytic models suggest that closing these schools had small positive effects on student learning and absenteeism. Students who attended chronically underperforming middle schools had lower ELA and mathematics gains and higher rates of absences compared to students who would have enrolled in the same school had it not closed. With two exceptions, this finding is consistent across all three years of closure, although the effects vary by grade, cohort and subject.

These findings must be considered in light of a few limitations. The analyses are based on post-hoc, observational data, and although I was able to closely match treatment

and control groups along a number of student characteristics, it is possible that the analyses remain biased by unobserved confounding characteristics that are related to both the treatment assignment and student academic growth. However, by accounting for student's elementary school in the matching process—in addition to multiple years of test score and attendance data—the models are relatively robust, given that the factors that influence elementary school choice are likely to be quite similar to the factors that influence middle school enrollment choices (and arguably student academic development).

Additionally, although the models may suffer bias stemming from the use two separate time periods, I was not able to control for any year effects. Specifically, it is possible that the district implemented an initiative that affected only students attending low-performing schools during one cohort's middle school years and not during the comparison cohort school years. However, the consistent nature of my findings across all three years of closure—and the fact that across years of closure the control and treatment groups share some common school years—provides some degree of confidence about my findings.

Robustness Check: Fixed-Effects Analyses

As a robustness check on these results and to address the potential biases discussed above, I use student fixed-effects models to examine the impact of closure, an approach that allows me to hold constant implicitly both observed and unobserved time-invariant student characteristics. This second approach also allows me to control for year effects, a second potential bias in my primary models. This analysis is slightly different than the above difference-in-differences model. First, it measures the impact of attending

a chronically failing school compared to the counterfactual of attending any other *current* alternative receiving school. Second, it examines the impact of attending a to-be-closed school at any point during middle school, rather than simply enrolling in a to-be-closed school in sixth grade.

Table 5-3 summarizes the within-student results for ELA and mathematics learning, which describe the impact of attending a chronically underperforming school. My results suggest a small negative impact of attending a to-be-closed school on both ELA and mathematics learning, and further, that the negative impact increases very slightly by length of time enrolled in a chronically underperforming to-be-closed school, as indicated by the closure prior exposure measure and exposure-by-attendance-at-a-closed-school interaction coefficients. On average, the effect of attending a closed schools is -0.03 SD, and this effect decreases by -0.08 SD for each year enrolled in a closed school. Results for mathematics are of the same magnitude.

Table 5-3
Within-Student Estimates of Attending a School Identified for Closure on ELA and Mathematics Learning (n= 9,591 students)

	ELA	Mathematics
Intercept	-0.543***	-0.535***
	(0.008)	(0.008)
Enrolled in a to-be-closed school	-0.080***	-0.080***
	(0.014)	(0.014)
Exposure to to-be-closed school(number of prior years enrolled in to-be-closed school)	-0.083**	-0.081***
	(0.022)	(0.020)
Enrolled in*exposure to to-be-closed school	0.054*	0.051**
	(0.019)	(0.016)
Sixth grade	0.039**	0.050***
	(0.013)	(0.010)
Seventh grade	0.055*	0.056**
	(0.020)	(0.016)
Eighth grade	0.093**	0.064*
	(0.027)	(0.025)

*p<.05; **p<.01; ***p<.001. Observations are annual.

Figures 5-4 and 5-5 provide graphic displays of the estimates in Table 5-3, organized by subject. To simplify the presentation, the predictions assume that the one year of exposure to a closed school occurred in 6th grade rather than 7th or 8th grade, and that the two years of exposure occurred in 6th and 7th grade. The y-axis indicates achievement levels in standard deviation units. For both ELA and mathematics, the models predict an average gap of .17 SD by eighth grade between students who had not enrolled in a chronically underperforming school and students who attended a chronically underperforming school all three years of middle school. With only attending sixth grade at a to-be-closed school, we would predict a slightly smaller ELA and mathematics test score gap in eighth grade of .11 SD. I also examine impact of closure on student attendance rates, and find no impact of attending a chronically underperforming school on rates of student absences (See Table 5-4).

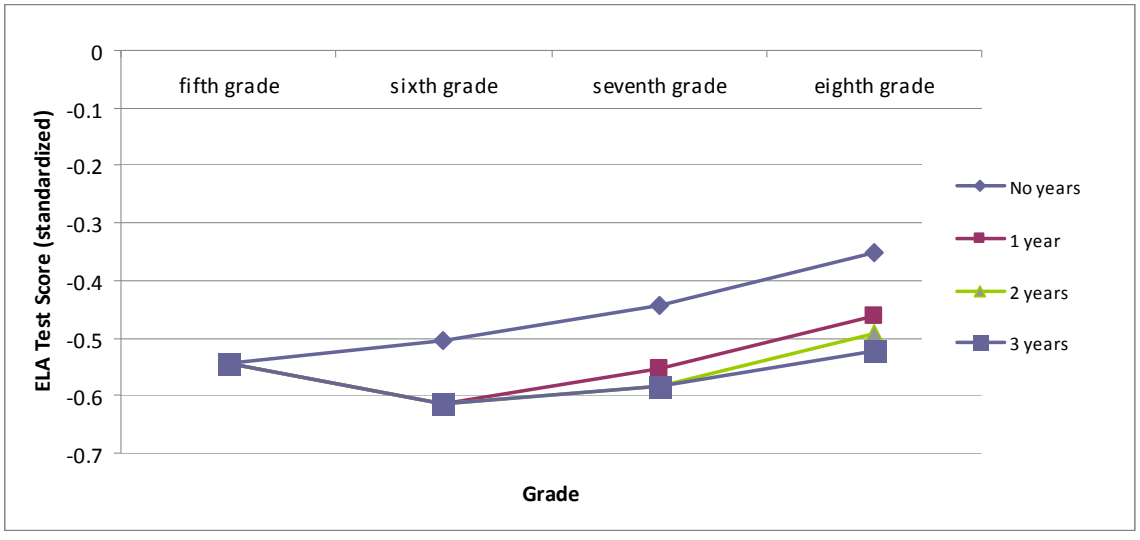


Figure 5-4. Predicted ELA test scores by years of exposure to a to-be-closed school.

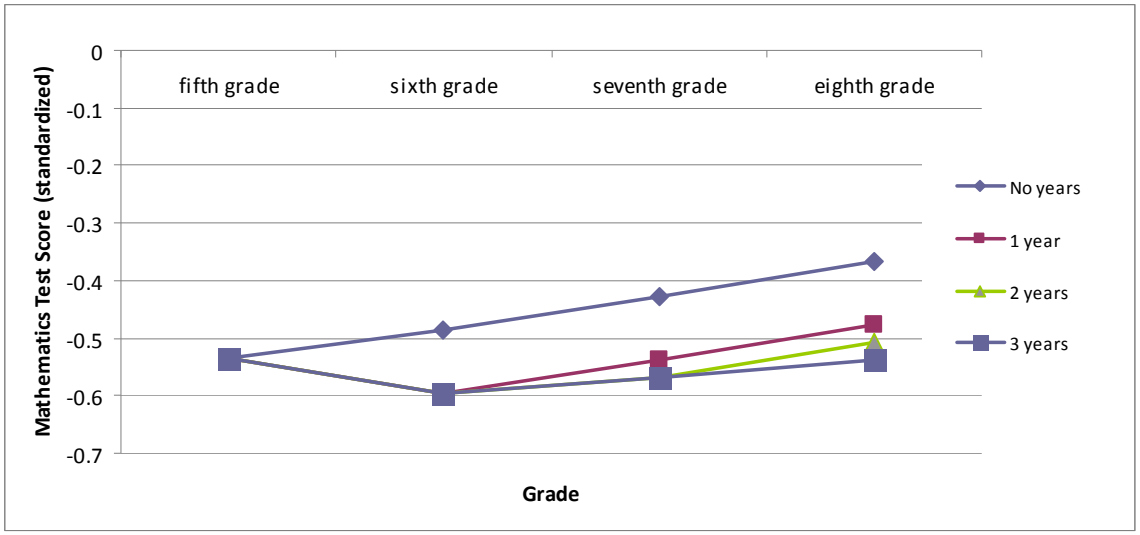


Figure 5-5. Predicted mathematics test scores by years of exposure to a to-be-closed school.

Table 5-4
Within-Student Estimates of Attending a School Identified for Closure on Absences
(n= 9,591 students)

	Absences (logged)
Intercept	2.661***
	(0.007)
Enrolled in a to-be-closed school	-0.027
	(0.014)
Exposure to to-be-closed school(number of prior years enrolled in to-be-closed school)	0.002
	(0.020)
Enrolled in*exposure to to-be-closed school	0.009
	(0.017)
Sixth grade	0.065***
	(0.011)
Seventh grade	0.178***
	(0.017)
Eighth grade	0.326***
	(0.025)

*p<.05; **p<.01; ***p<.001. Observations are annual.

Fixed-Effects Discussion

My fixed-effects results suggest a small, negative impact of attending a chronically underperforming to-be-closed school on both mathematics and ELA test

scores within individual students. Students who attended chronically underperforming middle schools learned less than they would have had they attended other alternative schools. The magnitude of the impact of closure on student test scores using a fixed-effect approach is strikingly similar to the magnitude of the findings using the matching difference-in-differences model, despite the slightly different populations and treatment of interest. The consistency of the results across the two models provide some additional confidence in the conclusion that the impact of closure on student learning is positive but quite small.

I did not find any impact on rates of student absences, a finding in contrast to that of my primary analyses using a difference-in-differences approach, in which attending a to-be-closed school appeared to increase student rates of absences. This discrepancy is likely in part due to the fact that the fixed-effects findings pertain only to the already more mobile students—those who switch schools during middle school—who are also likely to have high rates of absences, perhaps regardless of what school they attend. Nonetheless, the lack of impact of closure on attendance rates in this second model suggest that further research is needed on other student outcomes beyond test scores.

Additionally, one shortcoming of this fixed-effect strategy is that it does not control for time-varying unobserved characteristics that might change across middle school and be related to student learning and attendance outcomes. For example, students may switch middle schools because a sudden change in family income that necessitates moving apartments, and a sudden drop in family income is also likely to negatively affect student learning and attendance, negatively biasing the results. It seems possible, however, that students who move *in* to to-be-closed schools are just as

likely as students who move *out* of to-be-closed schools to be experiencing these kinds of disruptive conditions, a possibility that reduces concerns about this type of bias.

Chapter 6: Conclusion

Reformers and policy makers often lament the seemingly intractable nature of the practice of education—what Tyack and Cuban (1995) term the "grammar of schooling." Despite advances in knowledge and technology, instruction and the organization of schools have not followed pace. This would not be so troubling were all schools of equally high quality. Unfortunately, the schools least likely to change often are those who most need to change, and that serve the most disadvantaged clientele. Arne Duncan's comments on the new federal School Improvement Grant program typify this perspective: "For too long, educators have tinkered around the edges in low-performing schools, consigning generations of students of color to receiving an inferior education. It's time to transform chronically low-performing schools. It's time to put an end to the tireless tinkering" (USDOE Press Office, 2012).

An increasingly popular response to such calls for bolder, more dramatic education reform is the closure of chronically failing schools. School closures in several cities, including New York and Chicago, have accelerated over the last few years, partly in response to federal and state policies designed to foster this acceleration. The focus on closing schools rather than attempting to transform struggling schools represents a shift from a "schools" approach to a broader and more systemic notion of education reform, one that essentially gives up on the lowest-performing schools for the greater good of the system as a whole.

Despite the increased use of school closures as a reform, very little research has examined the effect of closure on student outcomes. This dissertation is an attempt to fill

this gap in the literature by examining the implementation of school closure in New York City over the past decade and its impact on students.

Summary of Findings

In my first results chapter, I describe the types of schools that New York City selects for closure. I find that schools that are closed evidence significantly lower school-average state test score exams and lower attendance rates compared to other middle schools, usually for several years prior to closure. The little variation that exists appears to be explained by the closing of a school for fiscal considerations rather than accountability purposes. They also tend to experience a small drop in enrollment prior to selection to closure. Moreover, in many ways the closed schools mirror the stereotypes of urban failing schools—students who attend these schools are almost exclusively Hispanic and Black, they are more likely to come from low-income families, and are more mobile than other middle schools students in the district. They are less likely, however, to come from non-English-speaking households. I also find that the students who enroll in these chronically under-performing middle schools do so already at a significant academic disadvantage; prior to enrolling in the to-be-closed middle schools, entering students have significantly lower test scores and higher rates of absences and tardies than students in other middle schools. Additionally, school-average test scores and absences of entering students appear to drop slightly just prior to a schools' selection for closure. These findings suggest that in part, closed schools perform poorly because their entering students are some of the most vulnerable and high-needs students in the District.

I also examine characteristics of the reconstituted schools that replace the closed schools, and find that reconstituted schools enroll higher performing students with fewer absences and tardies in the year prior to enrolling, compared to students who enroll in the schools that are later closed. However, with few exceptions, reconstituted schools appear to be serving similar types of students to those served by the closed schools in terms of race/ethnicity and home language, mobility and eligibility for free/reduced lunch. This finding provides some evidence that the district was not particularly successful in altering the concentrations of poor students within the schools, one of the stated goals of the policy. In contrast, the reconstituted schools tend to be less *academically* segregated than their closed counterparts.

In the second results chapter, I examine the District's implementation of the phase-out process, focusing on student mobility in the two years prior to closure. I find some small variation across District administrations in student mobility. Specifically, schools phased-out prior to 2002 tend to have slightly higher rates of mobility, while mobility in schools phasing-out after this year are similar to other District schools that are not phasing-out. Additionally, with the exception of two school years, my analysis suggests neighboring schools are not burdened by the enrollment of high numbers of students fleeing the phase-out process. Students that do elect to leave the schools during phase-out tend to have slightly higher test scores than those who stay, but higher rates of absences and tardiness.

In the third results chapter, I examine the impact of closure on student outcomes using a matching difference-in-differences approach, comparing student test scores and rates of absences between students who attended to-be-closed schools with students who

would have attended the schools had they not closed. I focus this analysis across three cohorts of schools that were closed in 2006, 2007 and 2008. I find across all three cohorts a small, positive effect of school closure on student test scores. Students who attended chronically underperforming middle schools had lower ELA and mathematics gains compared to students who would have enrolled in the same school had it not closed. Similarly, I find a small positive effect of closure on student absences: students who attend to-be-closed schools have slightly higher rates of absences compared to their counterparts who attended alternate schools.

Finally, as a robustness check for the results obtained via the matching, difference-in-differences models, I employ a second set of analyses to examine the impact of closure. Specifically, I use a student fixed-effects approach to compare student test-score and attendance trajectories within individuals before and during middle school. The results of these models are similar to those produced by my main models: students learn slightly less at chronically underperforming schools, compared to what would have happened had they attended an alternate school. However, in contrast to my difference-in-differences model results, I did not find an effect on student absences. This differential finding warrants further study, but is possibly attributable to the fact that the fixed-effects results pertain only to mobile students—those who move middle schools during the middle grades.

Limitations

External Validity and Generalizability

There are a number of limitations to my findings, some of which relate to the extent to which the results can be generalized to other contexts. The data are from one

city that arguably is an anomalous and unique context. New York City is the largest school district in the U.S., and its sheer size, number of schools and perhaps most importantly, density of schools might prevent generalizing these findings to other urban contexts. Because schools are so geographically close, school closure in New York City might be less disruptive to students than in other areas. The labor market in New York is also somewhat unique in its size and relatively high level of education, meaning that reconstituted schools may have an easier time hiring replacement teachers than reconstituted schools in other, less populous urban areas. For example, while implementing school turnarounds under the SIG program, Prince George's County Public Schools, a diverse Maryland district in suburban Washington, D.C., was forced to abandon closure and reconstitution and use a school reform option that did not require replacing the staff because of severe difficulties in hiring new teachers. Not unexpectedly, rural areas also have faced difficulties in re-staffing schools under the SIG program (Scott et al., 2012).

A second potential limitation related to the generalizability of my findings is the small number of closed schools I include in the impact study. Having so few schools in the study sample provides us with less certainty of whether the closed schools are representative of the future population of closed schools in the District to which we would like to generalize the findings, much less school closures outside of the District. And even if this small sample of schools closed are representative of closed middle schools more widely, the small sample size hinders the detection of statistical significance. However, this small sample-size problem is difficult to avoid given that the

phenomenon—closing schools—is not a common occurrence in any given year among schools with these particular grade spans.

A third concern regarding the external validity of my results is related to the sample of my primary analysis, the matched students. The difference-in-difference analysis focuses on examining the impact of closure for a unique population of students, those who attend closed schools or would have done so. Extrapolating these findings to other students in other districts requires assuming that attendance patterns at schools later slated for closure are similar to those in New York City. Additionally, the lack of variability in student outcomes in closed schools and their matched sample complicates the detection of effects.

Questions for Future Research

My findings provide some preliminary evidence about the impact of closure, but do not go far enough in examining the second phase of closure; reconstitution. Specifically, my research examines students outcomes across all control scenarios, regardless of the type of school in which the control students enrolled. Further research is needed to examine whether the treatment effect of attending a chronically failing school differs based on the counterfactual state—that is, whether the effect of closure differs by whether a control student attended the reconstituted school versus another neighboring middle school. This is a particularly important question to examine given that reconstitution is the counterpart to school closure, and the findings in Chicago that effects varied by the type of school students moved to following closure.

Another question is whether the findings vary by school characteristics, such as student body and staff composition and size of the alternate school and whether the

reconstituted school is a wall-to-wall middle school, a small school sharing a space with multiple schools, or has a particular program of focus. Additionally, it seems likely that the effects of the treatment might differ for different types of students. For example, relatively higher-performing students might perform just as well in either condition, while the lowest-performing students might be more responsive to different types of support and thus experience greater gains following school closure.

Additionally, my primary analyses focused on the impact of enrolling in a closed school regardless of whether the students remained for the full three years of middle school. Although measuring the impact of enrollment is an important first step because it allows me to account for any impacts of attending a failing school on mobility and potential negative effects on student learning, I may have underestimated the effect of *attending* the closed school. My analyses average effects across students who leave the to-be-closed schools after the first or second year and those who remain at the to-be-closed school for all three years. Future research should examine treatment dosage and tease out the impact of attending a closed school over one, two or three years compared to attendance at alternate middle schools.

My primary examination of impacts is limited to outcomes following the first few years of closure, in which the reconstituted schools are still quite young. Future research should examine student outcomes as the reconstituted schools develop, as some research on the establishment of new schools suggests that these schools have higher impacts after the first few years of establishment (Bloom & Unterman, 2012).

Additionally, my dissertation focuses on the impact of school closure on students who would have attended the closed school—a logical next step would be to focus on

reconstitution and specifically, on all students who attend reconstituted schools, not just those who would have attended the closed schools. Another treatment group of interest is the students who would have attended the existing neighboring schools regardless of school closure. De la Torre and Gwynne (2009) note that Chicago schools that accommodate an influx of students from closed schools, most of whom are low-achieving, are stretched to such an extent that they are unable to provide the same level of prior services. That study's descriptive analyses indicate that these displaced students end up moving to a small handful of schools that tend to be in close proximity to the closed schools, despite the fact that students had the option to enroll in any school with available space. My description of the phase-out process suggests that patterns in New York City might be different than in Chicago, but it seems important to examine the impact of the school closure policy not just on students who would have attended the closed school had it not been closed or who attend the newly reconstituted school, but on the students who attend the neighboring receiving schools after the school closes. If these receiving schools experience an influx of low-performing students, they may struggle to maintain the existing quality of instruction, in which case we might expect student achievement to remain constant or to decrease in comparison to the achievement levels that might have resulted with attendance at the chronically failing schools.

Finally, future research needs to examine the benefits and costs of closure for teachers and neighborhoods where closure occurs. The District administration rests its main argument for closure on the presumed benefits for students, leaving the impression that any negative implications of closure for teachers and neighborhoods, such as decreased neighborhood social capital, are problems that should not be factored in when

examining costs and benefits of school closure. However, the relatively minor benefits of closure for students that potentially can be achieved through other less disruptive reforms call for a closer examination of the potentially large costs of closure for other non-student stakeholders.

Implications

Despite these limitations, my findings suggest some implications for policy. Specifically, school closure appears to be a somewhat effective policy to improve student academic outcomes. It is not clear, however, whether the policy is *efficient*. That is, it is important to note that the gains of students who attend alternate schools are only slightly larger than the gains of students attending schools slated to be closed—less than 0.2 SD over three years. For comparison, a meta-analysis of comprehensive school reform suggests that these types of reforms result in effects of .16 SD on average (Borman et al. 2003). In contrast, meta-analyses of reforms that target specific instructional practices in middle and high school appear to be slightly larger, although variable depending on the focus. While curricula program effects were small with average effect sizes of .10 SD in mathematics, programs that focused on instructional methods using extensive professional development resulted in significantly larger effects, with a median of 0.34 SD in mathematics and .021 in literacy (Slavin, Lake & Groff, 2008). And recent research on school turnaround models that replace leadership but not the full staff, or that replace leadership and some staff but not students—both potentially less politically-fraught policy alternatives to closure, for example—finds effects to be of similar magnitude as my findings on the impact of closure (de la Torre et al., 2012). Given the significant amount of resources expended by the District as it pushes ahead for closures

and by school and community members as they oppose school closures, the potential negative effects of closure on teachers and neighborhoods, the acrimonious public debates that have in some years been accompanied by lawsuits, and the substantial efforts required to close a school and open new ones, the cost of school closures may be too high in light of the small effects suggested here. More research is needed, therefore, that compares impacts of closure to other types of contemporary school turnaround reforms that are less divisive and disruptive to local communities and families, and to the staff that work in schools that are closed. And to fully understand the implications of closure, future research is required to identify the fiscal costs associated with closure, compared to the costs of other policies with effects that are of comparable size.

The central finding from my dissertation provides a small step forward in sorting out the rival theories that support or rebut the argument for the use of school closure. Specifically, the small positive effect of school closure suggests that human capital theory cannot be outright rejected. It remains possible that removing ineffective teachers is the most effective manner to improve student learning at chronically underperforming schools. However, more evidence is required on the possible detrimental effects of school closure, such as whether the teachers at the closed schools are able to be more effective in other schools or with other students. And, as mentioned earlier, research comparing the effects of school closure to specific alternative school reform options is needed. For example, to understand whether the improvement in student outcomes is due to transformations to school-based social capital or human capital reasons would require a comparison of school closure effects to less disruptive school turnaround reforms, particularly given the similar effect sizes found in the most recent research in Chicago.

Research on the use of more targeted professional development is also required to address the question of whether teachers at the chronically underperforming schools are truly unable to change their instruction with the right supports.

This dissertation also does little to resolve one of the more troubling dilemmas of school closure. The District's desire to lessen school segregation conflicts with the benefits that family and community social capital might afford students, and is in many respects at odds with the school community's desire to ensure that the same students and families will be served by their neighborhood school after a school is closed. Perhaps the fact that the District has managed to establish reconstituted schools that serve students who are similar demographically but less segregated in terms of academic achievement—arguably the most pernicious type of school segregation—is the best middle ground.

Closure in Context

Closing schools for poor student performance represents a new approach to reform. Indeed, it is not school reform per se, but rather, a broader education reform in which the targets of reform are educational systems, rather than the processes and practices found within individual schools. One result is that the school-closure approach to reform does not prescribe specific methods—it is another in a long line of reforms that do not directly address teaching and learning and the processes that encompass classroom practice. Rather, similar to many contemporary school reforms, such as those related to school choice, decentralization and school governance, charter schools, class and school size, use of technology, school funding, principal autonomy and merit pay, school closure targets school structures and not instruction. These structural reforms are a means to an end—mediating strategies that are only effective in their ability to facilitate or create the

conditions for more effective instruction. They are themselves not able to directly touch the work that occurs inside the classroom.

Proponents of closure argue that instruction at these chronically underperforming schools simply can't be reformed. However, there is little evidence that these schools have implemented targeted, evidenced-based *instructional* programs. Indeed, the New York City DOE, despite the scale and number of education reforms implemented over the past decade, has been remarkably agnostic on instructional practices and content (O'Day & Bitter, 2011). This agnosticism is likely due in part to the fact that closure and other structural reforms are easier to understand, easier to implement and easier to measure than the more complicated classroom interactions that take place between students and teachers. However, the District has also purposefully backed away from instituting policies to manage instruction in any systemic manner, predicated on the idea that autonomy in exchange for increased accountability is the most effective means to increase student achievement (O'Day & Bitter, 2011). While it may be that allowing schools the autonomy to select their own instructional approaches has significant benefits for some types of schools that are already functioning at some baseline level, these chronically underperforming schools may be less likely to benefit from the autonomy that this approach to reform is designed to allow. It also seems likely, given the years of languishing on city, state and federal accountability "worst school" lists prior to closure, that the problem with these schools is not the lack of incentives to improve. Moreover, with the advent of teacher-value-added policies and techniques, incentives to improve student learning are already on the rise.

Rather, before the lowest-performing schools are slated for closure, the District should consider taking the shortest route to effecting instructional change by directly assessing and targeting capacity at the instructional core, with focused attention to the evidence base on how and what to teach. Despite education research's reputation as weak and often unsubstantiated, there is in fact a significant amount of research that has identified specific instructional practices and content associated with large impacts on student learning, even taking into account school conditions and resources. Even if, as some proponents of closure claim, teacher capacity is a fixed attribute, the district can focus on increasing capacity of the other elements of instruction by aligning the content and delivery of instruction to what we know is effective for student engagement and learning. This type of approach requires a shift in focus from autonomy to one of capacity building, which would require sustained support and guidance based on knowledge about good instruction to allow for incremental sustainable change. For example, the district's current recommended middle school mathematics curriculum, Glencoe's Impact Mathematics, has no evidence of effectiveness (Slavin, 2009). One initial step could be to require struggling schools to select from curricula that have evidence of effectiveness. And perhaps the newly-introduced classroom observations tools, which are connected to what we know about effective instructional practices, are a promising method to link instructional practices more tightly to student outcomes—though it seems more likely that coupling the observations with accountability press may not allow for honest appraisals of practice. Extensive guidance across the District, targeting the instructional core, coupled with knowledge about what and how to teach that is relevant to the conditions these schools face could serve to improve student

learning before school closure is necessary. Given the relatively small benefits of school closure reported in my dissertation and the potentially disruptive nature of the reform, it seems prudent that school districts explore alternative avenues for their efforts to improve struggling schools.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95-135.

Advocates for Children of New York (ACNY) and the Asian American Legal Defense and Education Fund (AALDEF). (2009 June). *Empty promises.* New York: Authors.

Ahn, T. & Vigdor, J. (2009 October). *Does No Child Left Behind have teeth? Examining the impact of federal accountability sanctions in North Carolina.* Paper presented at Institute for Federalism and Intergovernmental Relations Martin School Workshop.

Aladjem, D. K., Birman, B. F., Orland, M., Harr, J., Heredia, A., Parrish, T. B., et al. (2010). *Achieving dramatic school improvement: An exploratory study.* Washington, DC: US Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service.

American Institutes for Research [AIR] & SRI International [SRI]. (2005). *Getting to results: Student outcomes in new and redesigned high schools.* Washington DC: Authors.

Angrist, J. D., & Krueger, A. B. (1999). Empirical strategies in labor economics. *Handbook of labor economics, 3*, 1277-1366.

Arkin, M., & Kowal, J. (2005). *School Restructuring Options Under No Child Left Behind: What Works When? Reopening as a Charter School*. Naperville, IL: Learning Point Associates.

Austin, P. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational methods. *Multivariate Behavior Research , 46*(3), 399-424.

Baker, A. (2012 June 29). Arbitrator halts City's plan to overhaul 24 schools. *New York Times.* Retrieved from http://www.nytimes.com/schoolbook/2012/06/29/city-loses-arbitration-on-staffing-for-24-turnaround-schools/

Ballou, D. & Springer, M.G. (2009). *Achievement trade-offs and No Child Left Behind.* Working Paper. Washington, DC: Urban Institute.

Barber, B. (1961). Social-class differences in educational life-chances. *Teachers College Record, 63*, 102-13.

Barrow, L., Claessens, A., & Schanzenbach, D.L. (2010 February). *The impact of small schools in Chicago.* Retrieved from: http://econweb.tamu.edu/common/files/workshops/PERC%20Applied%20Microeconomics/2010_2_10_Diane_Schanzenbach.pdf

Bates, D., Maechler, M. & Dai, B. (2011). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-42. Retrieved from http://cran.r-project.org/web/packages/lme4/index.html

Beale, S. (2011 March 2). Which Providence schools might close? *GoLocalProv News. Retrieved from:* http://www.golocalprov.com/news/exclusive-new-details-on-school-closings/

Bellandi, A. (2009, October 7). Chicago school reform blamed for teen deaths. *Education Week.*

Bloom, H. & Unterman, R. (2012 January). *Sustained positive effects on graduation rates produced by New York City's small public high schools of choice.* New York: MDRC Policy Brief.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*(2), 125.

Borman, G., & Dowling, M. (2010). Schools and inequality: A multilevel analysis of Coleman's Equality of Educational Opportunity data. *The Teachers College Record, 112*(5), 1-2.

Boyd, D., Grossman, P., Lankford, H., Loeb, S. & Wyckoff, J. (2008 May). *Who leaves? Teacher attrition and student achievement.* (NBER Working paper No. 14022).

Brady, R. C. (2003). *Can failing schools be fixed?* Washington, D.C.: Thomas B. Fordham Foundation.

Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (2009). *Estimating principal effectiveness* (Working Paper 32). Washington, D.C.: The National Center for Analysis of Longitudinal Data in Education Research, The Urban Institute.

Burgess, S., Propper, C., Slater, H., & Wilson, D. (2005). *Who wins and who loses from school accountability?: The distribution of educational gain in English secondary schools* (CEPR Discussion Paper No. 5248*).*

Carless, W. (2012 July 12). School closure plan killed before it starts. *Voice of San Diego.* Retrieved from

http://www.voiceofsandiego.org/education/article_b92016b6-caf1-11e1-b1e0-001a4bcf887a.html.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, *26*, 499–531.

Center on Education Policy (2010 August) *School Improvement Grants present uncertainty and opportunity.* Washington, DC: Author.

Christ, L. (2010 December 7). DOE Adds 14 More Schools to Phase-out List Bringing Total to 26. *NY1*. Retrieved from: http://www.ny1.com/content/news_beats/education/130195/doe-adds-14-more-schools-to-phase-out-list-bringing-total-to-26/

Christ, L. (2011 February 1). Two dozen protestors arrested during rally against school closures. *NY1*. Retrieved from: http://www.ny1.com/content/top_stories/133096/two-dozen-protesters-arrested-during-rally-against-school-closures

Clark, D., Martorell, P., & Rockoff, J. (2009). *School principals and school performance* (CALDER working paper no. 38). Washington D.C.: Urban Institute.

Cohen, D. (2007). Conclusion: Problems in education policy and research, in Furhman, Cohen & Mosher (Eds.) *The state of education policy research* (pp. 349-371). Mahwah, NJ: Lawrence Erlbaum Associates.

Coleman, J. (1988). Social capital in the creation of human capital. *The American Journal of Sociology*, 94 (S95-S120).

Cramer, P. (2008a Dec 5). UWS parents gear up for renewed diversity fight over school closure. *Gotham Schools.* Retrieved from http://gothamschools.org/2008/12/05/uws-parents-gear-up-for-renewed-diversity-fight-over-school-closure/.

Cramer, P. (2008b Dec 9). Elected parents learn of school closure by e-mail. *Gotham Schools.* Retrieved from http://gothamschools.org/2008/12/09/elected-parent-leaders-learned-of-school-closure-by-e-mail/.

Cramer, P. (2011a July 21). Judge rejects UFT-NAACP claims, allows co-locations, closures. *Gotham Schools.* Retrieved from: http://gothamschools.org/2011/07/21/judge-rejects-uft-naacp-claims-allows-co-locations-closures/

Cramer, P. (2011b Jan 31). As closure votes near, thoughts on what will follow for students. *Gotham Schools.* Retrieved from: http://gothamschools.org/2011/01/31/as-closure-votes-near-thoughts-on-what-will-follow-for-students/

CTB/McGraw Hill (2006 December). *New York State Testing Program 2006: Technical report.* Monterey, CA: Author.

Cuban, L. & Tyack, D. (1995). *Tinkering toward utopia: A century of public school reform.* Cambridge, MA: Harvard University Press.

Darling-Hammond, L. (2000). Futures of teaching in American education. *Journal of Educational Change, 1*(4), 353-373.

Dawsey, C. (2012 July 18). Feds evaluating complaints made about Detroit Public Schools closures. *Detroit Free Press.* Retrieved from

http://www.freep.com/article/20120718/NEWS01/207180343/Feds-evaluating-complaints-made-about-Detroit-Public-Schools-closures

de la Torre, M., & Gwynne, J. (2009). *When schools close: Effects on displaced students in Chicago Public Schools*. Chicago, IL: Consortium on Chicago School Research.

de la Torre, M., Allensworth, E., Jagesic, S., Sebastian, J., Salmonowicz, M., Meyers, C., & Gerdeman, D. (2012 Feb). *Turning around low-performing schools in Chicago.* Chicago, IL: Consortium on Chicago School Research.

Deere, D. Strayer, W. 2001. Putting schools to the test: school accountability, incentives and behavior. Working paper, Department of Economics, Texas A&M University.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181.

Dika S. & Singh, K. (2002). Applications of social capital in educational literature: A critical synthesis. *Review of Educational Research*, *72*(1), 31-60.

Elmore, D. (2006 November). *The problem of capacity in the (re)design of educational accountability systems.* Paper presented at Campaign for Educational Equity conference "Examining America's Commitment to Closing Achievement Gaps" held at Teachers College, Columbia University. *Retrieved from:* http://devweb.tc.columbia.edu/manager/symposium/Files/95_ElmoreCapacityPaper_10-5.pdf.

Epstein, J & Sanders, M. (2002). Family, school, and community partnerships. In: Bornstein, M. (Ed.) *Handbook of parenting: Vol. 5. Practical issues in parenting*. Erlbaum; Mahwah, NJ: 507-437.

Fertig, B. (January 25 ,2011). At schools that closed, more students graduate and more drop-out. *WNYC.* Retreived from: http://www.wnyc.org/articles/wnyc-news/2011/jan/25/schools-closed-more-students-graduate-and-more-drop-out/

Figlio, D. & Rouse, C. (2006). Do accountability and voucher threats improve low-performing schools? *Journal of Public Economics 90*, 239-255.

Fine, M. (1993). [Ap]parent Involvement: Reflections on Parents, Power, and Urban Public Schools. & Responses. *The Teachers College Record, 94*(4), 682-729.

Finn, C. (1987) Education that works: Make the schools complete. *Harvard Business Review, 65,* 64.

Freedman, S. (2008 May 7). Failings of one Brooklyn schools may threaten a neighbor's success. *New York Times.* Retrieved from: http://www.nytimes.com/2008/05/07/education/07education.html?_r=1&scp=1&sq=Lafayette%20Dewey%20Freedman&st=cse

Fung, A. (2004) *Empowered participation: Reinventing urban democracy*. Princeton, NJ: Princeton University Press.

Galletta, A. and J. Ayala (2008). Erasure and survival: Creating a future and managing a past in a restructuring high school. *The Teachers College Record, 110*(9), 1959-1985.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*, 915–945.

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University Press.

Gewertz, C. (2006 July 11). Federal officials to probe complaint against NYC small high schools. *Education Week.* Retrieved from: http://www.edweek.org/ew/articles/2006/07/12/42brief-3.h25.html?r=16207281

Gittrich, G. (2001 Sept 2). Changes To Battle Confusion. *New York Daily News.* Retrieved from http://articles.nydailynews.com/2001-09-02/news/18364417_1_middle-schools-high-school-reforms-special-education-reforms/2

Glazerman, S., Dolfin, S., Bleeker, M., Johnson, A., Isenberg, E., Lugo-Gil, J., & Ali, M. (2008). *Impacts of comprehensive teacher induction: Results from the first year of a randomized controlled study.* Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Goldstein, J., Kelemen, M. and Koski, W. (1998). *Reconstitution in Theory and Practice: The Experience of San Francisco.* Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).

Gootman, E.(2006 June 16).City's new small schools are focus in inquiry. *New York Times.* Retrieved from: http://www.nytimes.com/2006/06/16/nyregion/16schools.html?_r=1

Gottfried, M. (2009). Excused versus unexcused: How student absences in elementary school affect academic achievement. *Educational Evaluation and Policy Analysis 31*(4), 215–229.

Grady, M. & Beretvas, S. (2010). Incorporating student mobility in achievement growth modeling: A cross-classified multiple membership growth curve model. *Multivariate Behavioral Research*, *45*, 393-419.

Graham, J.W. (2008). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*(6), 1-28.

Grolnick, W., & Slowiaczek, M. (1994). Parents' involvement in children's schooling: A multidimensional conceptualization and motivational model. *Child Development 65,* 237–252.

Halaby, C. N. (2004). Panel models in sociological research: Theory into practice. *Annual Review of Sociology, 30*, 507-544.

Hallinger, P., & Heck, R. H. (1996). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly, 32*(1), 5.

Hanushek, E.A., Kain, J.F., Markman, J.M. and Rivkin, S.G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics, 18,* 527-544.

Harder, V.S., Stuart, E.A., and Anthony, J. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods, 15*(3), 234-249.

Harries, G. (2009, September) *Testimony for the Aspen Institute NCLB Turnaround Committee.*

Hartocollis, A. & Holloway, L. (1999 June 22). In a big shake-up, Crew will close 13 failing schools. *New York Times.* Retrieved from http://www.nytimes.com/1999/06/22/nyregion/in-a-big-shake-up-crew-will-close-13-failing-schools.html

Hartocollis, A. (1997 October 26). Small-schools conference lands Crew in a dispute. *New York Times.* Retrieved from http://www.nytimes.com/1997/10/26/nyregion/small-schools-conference-lands-crew-in-a-dispute.html

Hassel, B. C., & Steiner, L. (2003). *Starting fresh: A new strategy for responding to chronically low performing schools*. Chapel Hill, NC: Public Impact, LLC.

Hastings, J. S., & Weinstein, J. M. (2008). Information, School Choice, and Academic Achievement: Evidence from Two Experiments. *Quarterly Journal of Economics, 123*(4), 1373-1414.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies, 64*(4), 605.

Helms, A. D. & Price, M. (2010 November 9). Board closes Waddell, saves Harding. *Charlotte Observer.* Retrieved from http://www.charlotteobserver.com/2010/11/09/1824769/more-than-100-will-address-cms.html

Hemphill, C., Nauer, K., Zelon, H., & Jacobs, T. (2009). *The new marketplace: How small-school reforms and school choice have reshaped New York City's high schools*. New York: The New School Center for New York City Affairs.

Henderson, A. T., & Mapp, K. L. (2002). *A new wave of evidence: The impact of school, family, and community connections on student achievement*. Austin, TX: National Center for Family & Community Connections with Schools, Southwest Educational Development Laboratory.

Henig, J., Gold, E., Orr, M., Silander, M. & Simon, E. (2011). Parent and Community Engagement and the Sustainability Challenge. In O'Day, Bitter & Gomez (Eds.) *Education Reform in New York City*. Cambridge, MA: Harvard Education Press.

Herlihy, C. M., & Kemple, J. J. (2004). *The Talent Development middle school model: Context, components, and initial impacts on students*. New York: MDRC.

Hernandez, J.C. (2009, February 2). To close a school: A decision rooted in data but colored by nuance. *The New York Times.* Retrieved from www.nytimes.com/2009/02/02/education/02closed.html.

Hernandez, J.C. (2011, February 4). Bloomberg calls behavior of Black's critics an embarrassment. The *New York Times*. Retrieved from http://cityroom.blogs.nytimes.com/2011/02/04/bloomberg-calls-behavior-of-blacks-critics-an-embarrassment/

Hess, F. & Gift, T. (2009 February). School turnarounds: Resisting the hype, giving them hope. *Education Outlook*, *2.* Washington, DC: American Enterprise Institute for Public Policy Research.

Hess, G. A. (2003). Reconstitution—three years later: Monitoring the effect of sanctions on Chicago high schools. *Education and Urban Society, 35*(3), 300-327.

Hill, H. C. (2007). Learning in the teaching workforce. *The Future of Children*, 17: 111–127.

Hill, J. (2008). Discussion of research using propensity score matching: Comments on "A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003" by Peter Austin. *Statistics in Medicine*, *27*, 2055–2061.

Hill, N. E., & Tyson, D. F. (2009). Parental involvement in middle school: A meta-analytic assessment of the strategies that promote achievement. *Developmental Psychology, 45*(3), 740.

Ho, D. E., Imai, K., King, G. and Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*, 199–236.

Ho, D.E., Imai, K., King, G., and Stuart, E.A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software.*

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*(396), 945-960.

Honaker, J., King, G. and Blackwell, M. (2011). *Amelia II: A program for missing data.* Retrieved from: http://gking.harvard.edu/amelia.

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher, 33*(8), 16.

Horng, E. L., Klasik, D., & Loeb, S. (2009). *Principal Time-use and School Effectiveness*. (Working Paper 34). Washington, DC: CALDER Urban Institute.

Ingersoll, R. M. (1999). The Problem of underqualified teachers in American secondary schools. *Educational Researcher, 28*(2), 26-37.

Inside schools (2011). *Types of schools: Zoned schools and other options.* New York: Author. Retrieved from: http://insideschools.org/?s=1&a=23#elem

Israel, G., Beaulieu, L. & Hartless, G. (2001). The influence of family and community social capital on educational achievement. *Rural Sociology, 66*(1), 43–68

Jacob, B. (2005). Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics, 89*(5-6), 761-796.

Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban Education, 40*(3), 237.

Jeynes, W. H. (2007). The relationship between parental involvement and urban secondary school student academic achievement. *Urban Education, 42*(1), 82.

Kandile, C.Y. (2011 March 3). Federal probe launched into BPS school closings. *The Bay State Banner*. Retrieved from: http://www.baystatebanner.com/local12-2011-03-03

Kane, T. J. and D. O. Staiger (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*,*16*(4), 91-114.

Kaye, R. (2010 February 24). All teachers fired at Rhode Island School. *CNN*. Retrieved from: http://articles.cnn.com/2010-02-24/us/rhode.island.teachers_1_teachers-union-troubled-school-reading-specialists?_s=PM:US

Kemple, J.J., and Hurlihy, C.M. (2004). *The Talent Development Middle School model.* New York: Manpower Demonstration Research Corporation.

Kilbride, K. (2012 June 29). South Bend school closure timeline revealed. *South Bend Tribune.* Retreived from http://www.southbendtribune.com/news/sbt-south-bend-school-closure-timeline-revealed-20120629,0,1733554.story

Kirshner, B., Gaertner, M., & Pozzoboni, K. Tracing transitions. *Educational Evaluation and Policy Analysis, 32*(3), 407-429.

Kotter, J. P. (1995). Leading change: Why transformation efforts fail. *Harvard Business Review, 73*(2), 19–27.

Kowal, J. M., & Hassel, E. A. (2005). *Turnarounds with new leaders and staff*. Naperville, IL:Learning Point Associates/North Central Regional Educational Laboratory (NCREL).

Ladd, H. (1999). The Dallas School accountability and incentive program: an evaluation of its impacts on student outcomes. *Economics of Education Review,* 18, 1 –16.

Ladd, H. (2009*). Teachers' perceptions of their working conditions: How predictive of policy-relevant outcomes?* (CALDER Working Paper 33). Washington, DC: The Urban Institute.

Lareau, A. (1987). Social class differences in family-school relationships: The importance of cultural capital. *Sociology of education, 60*(2), 73-85.

Lee, V. E., & Smith, J. B. (1996). Collective responsibility for learning and its effects on gains in achievement for early secondary school students. *American Journal of Education, 104,* 103-147.

Lee, V. and Smith, J.B. (1995). Effects of high school restructuring and size on gains in achievement and engagement for early secondary school students. *Sociology of Education, 68*(4)*,* 241-270.

Lee, V.E. and Ready, D.D. (2007). *Schools within schools: Possibilities and pitfalls of high school reform.* New York: Teachers College Press.

Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004*). Review of research: How leadership influences student learning.* New York: The Wallace Foundation.

Lisberg, A & Monahan, R. (2010, July 16). 19 schools city blocked from closing barely have enough kids enrolled to stay open. *New York Daily News.* Retrieved from: http://www.nydailynews.com/ny_local/education/2010/07/16/2010-07-16_19_city_schools_city_blocked_from_closing_barely_have_enough_kids_enrolled_to_st.html

LoGerfo, L., Nichols, A., & Reardon, S. F. (2006). *Achievement gains in elementary and high school*. Washington DC: The Urban Institute.

Loveless, T. (2010 January). *How well are American students learning?* (Vol. II, No. 4). Washington, DC: Brookings Institution Press.

Lumley T (2010*). mitools: Tools for multiple imputation of missing data.* R package version 2.0.1, retrieved from: http://CRAN.R-project.org/package=mitools.

Maxwell, L.A. (2010 June 15). Kentucky high school launches turnaround journey. *Education Week.* Retrieved from: http://www.edweek.org/ew/articles/2010/06/16/35kyturnaround_ep.h29.html?r=1866715968

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *The American Journal of sociology, 83*(2), 340-363.

Meyer, J. W., & Rowan, B. (1977). The structure of educational organizations. In J.W. Meyer (Ed.) *Environments and Organizations*. San Francisco, CA: Jossey-Bass: 78–109.

Neal, D. & Schanzenbach, D.W. (2010). Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics.* 92(2): 263-283.

New York 1 (2012 April 27). Public City school closings, overhauls for 2012. Retrieved from http://www.ny1.com/content/top_stories/154985/public-city-school-closings—overhauls-for-2012/#closure .

New York City Department of Education (NYCDOE) (2007 Summer). *New York City's English language learners.* New York: New York City Department of Education Office of English Language Learners. Retrieved from http://schools.nycenet.edu/offices/teachlearn/ell/DemoPerformanceFINAL_10_17.pdf

New York City Department of Education (NYCDOE) (2009). *Test Information.* Retrieved from: http://schools.nyc.gov/daa/test_info/default.asp#ELA.

New York City Independent Budget Office (IBO). (2010 January 25). *Comparisons between schools slated for closing and all other schools*. New York: Author. Retrieved from: http://www.ibo.nyc.ny.us/iboreports/schoolclosing012510.pdf

New York State Education Department (NYSED) (2010a December 9). *Education Commissioner announces 67 schools identified as "Persistently Lowest Achieving" with 58 of them also placed under registration review; Schools to be the focus of school improvement efforts.* Retrieved from: http://www.oms.nysed.gov/press/PersistentlyLowestAchievingAndSURRDec2010.html

New York State Education Department (NYSED) (2010b February 16). *Commissioner Steiner announces SURR schools list 14 schools removed; 4 schools added.* Retrieved from: http://www.oms.nysed.gov/press/SURR2010.html

New York State Education Department (NYSED) (2010c February). *Field guidance memorandum: Federal School Improvement Grants under section 1003(g) of the Elementary and Secondary Education Act (#2010/1).* Retrieved from: http://www.p12.nysed.gov/accountability/TitleI/sigfieldguidance.html

New York State Education Department (NYSED) (2011). *Percent distribution of public school students by racial/ethnic origin*. Retrieved from http://www.p12.nysed.gov/irs/statistics/public/2011/TABLE4.pdf

Normand S., Landrum M., Guadagnoli E., Ayanian J., Ryan T., Cleary P., McNeil B. (2001). Validating recommendations for coronary angiography following an acute myocardial infarction in the elderly: A matched analysis using propensity scores. *Journal of Clinical Epidemiology, 54,* 387–398.

O'Day, J. & Bitter, C. (2011). Improving instruction in New York City. In J. O'Day, C. Bitter & L. Gomez (Eds.), *Education Reform in New York City* (pp. 109-130). Cambridge, MA: Harvard Education Press.

O'Day, J., Bitter, C. & Talber, J. (2011). Introduction to the volume and Children First. In J. O'Day, C. Bitter & L. Gomez (Eds.), *Education Reform in New York City* (pp. 1-16). Cambridge, MA: Harvard Education Press.

Otterman, S. (2010a, Jan 21). City to get federal aid to fix or close weak schools. *New York Times*. Retrieved from http://cityroom.blogs.nytimes.com/2010/01/21/city-to-get-federal-aid-to-fix-or-close-weak-schools/#more-124377

Otterman, S. (2010b, Jan 24). Large high schools in the city are taking hard falls. *New York Times.* Retrieved from http://www.nytimes.com/2010/01/26/nyregion/26closings.html?_r=3&ref=nyregion

Otterman, S. (2010c, Jan 26). Themes from the tumult of school closings. *New York Times.* Retrieved from http://cityroom.blogs.nytimes.com/2010/01/26/themes-from-the-tumult-of-school-closings/

Otterman, S. (2011, February 3). Protesting school closings, a noisy annual ritual. *New York Times*. Retrieved from: http://www.nytimes.com/2011/02/04/nyregion/04panel.html

Oxley, D. (1989). Smaller is better. *American Education, 13,* 28-31, 51-52.

Penuel, W., Riel, M, Krause, A, & Frank, K. (2009). Analyzing teachers' professional interactions in a school as social capital: A social network approach. *Teachers College Record 111*(1), 124–163.

Phillips, A. (2011March 18). A Brooklyn school wins the right to stay open, but must shrink. *GothamSchools.* Retrieved from: http://gothamschools.org/2011/03/18/a-brooklyn-school-wins-the-right-to-stay-open-but-must-shrink/

Plasencia, P. & Agarwal, U. (2009, August 4). Un modelo que no sirve nuestros niños [A system that does not work for our children.] *El Diario.* Retrieved from http://www.impre.com/eldiariony/opinion/2009/8/4/un-modelo-que-no-sirve-nuestro-139534-1.html.

Pribesh, S., & Downey, D. B. (1999). Why are residential and school moves associated with poor school performance? *Demography, 36*(4), 521-534.

Raudenbush, S.W., and A.S. Bryk. *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage, 2002.

Raywid, M.A. (1995). *The subschools/small schools movement – Taking stock.* Madison, WI: Center on Organization and Restructuring of Schools.

Reback, R. (2008). Teaching to the rating: School accountability and the distribution of student achievement. *Journal of Public Economics,* 92, 1394-1415.

Reback, R., Rockoff, J., & Schwartz, H.L. (2011, January). *Under pressure: Job security, resource allocation and productivity in schools under NCLB* (NBER Working Paper No. 16745).

Rice, J. K. and B. Malen (2003). The human costs of education reform: The case of school reconstitution. *Educational Administration Quarterly, 39*(5), 635.

Rich, M. (2012 July 23). Enrollment off in big districts, forcing layoffs. *New York Times.* Retrieved from http://www.nytimes.com/2012/07/24/education/largest-school-districts-see-steady-drop-in-enrollment.html?_r=1&ref=education.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Robertson, J. (2010 March 10). Kansas City to shutter 26 schools. *Kansas City Star.* Retrieved from: http://www.mcclatchydc.com/2010/03/10/90191/kansas-city-to-shutter-26-schools.html

Robinson, G. (2010 Jan 25). Latest round of school closings sparks heated opposition. *Gotham Gazette.* Retrieved from: http://www.gothamgazette.com/article/education/20100125/6/3160

Robinson, G. (2011 January). With more school closings, debate about policy continues. *Gotham Gazette.* Retrieved from: http://www.gothamgazette.com/article/Education/20110131/6/3460

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*(2), 247-252.

Rockoff, J., and Turner, L. (2010). Short-run Impacts of Accountability on School Quality, *American Economic Journal, Economic Policy, 2*(4), 119-147.

Rondinone, D. (2010 Dec 7). DOE closing dozens of schools. *NY1*. Retrieved from: http://www.ny1.com/content/news_beats/education/130156/doe-closing-dozens-of-schools/

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika,, 70*(1), 41-55.

Rowland, C. (2011 Jan 31). Closing schools: Myth and mystery. *GothamSchools.* Retrieved from: http://gothamschools.org/2011/01/31/closing-schools-myth-and-mystery/

Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics, 29*, 159–184.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology, 66*(5), 688.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. Health *Services & Outcomes Research Methodology 2*, 169–188.

Rubin, D. B., & Thomas, N. (2000). Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates. *Journal of the American Statistical Association, 95*(450), 573-585.

Rubin, D.B. (1987) *Multiple imputation for nonresponse in surveys*. J. Wiley & Sons, New York.

Schwartz, A.E. & Steifel, L. (2006). Is there a nativity gap? Achievement of New York City elementary and middle school immigrant students. *Education Finance and Policy, 1*(1), 17-49.

Scott, C. (2008 September). *A call to restructure restructuring*. Washington DC: Center on Education Policy. Scott, C. (2008).

Scott, C., McMurrer, J., McIntosh, S., and Dibner, K.(2012 March). Opportunities and Obstacles: Implementing Stimulus-Funded School Improvement Grants in Maryland, Michigan, and Idaho. Washington DC: Center on Education Policy. Retrieved from http://www.cep-dc.org/displayDocument.cfm?DocumentID=397

Seltzer, M. H. (1995). Furthering our understanding of the effects of educational programs via a slopes-as-outcomes framework. *Educational Evaluation and Policy Analysis, 17*(3), 295.

Shin, P. (2000 Dec 13). Taken off review; 3 must close. *New York Daily News.* Retrieved from http://articles.nydailynews.com/2000-12-13/news/18160479_1_middle-schools-exams-education-commissioner-richard-mills

Singer, J.D., and J.B. Willett. *Applied longitudinal data analysis: Modeling change and event occurrence.* New York: Oxford University Press, 2003.

Slavin, R., Lake, C., & Groff, C. (2008). *Effective programs in middle and high school mathematics: A best-evidence synthesis.* Baltimore, MD: Center for Research and Reform in Education, Johns Hopkins University.

Smith, S. S., & Mickelson, R. A. (2000). All that glitters is not gold: School reform in Charlotte-Mecklenburg. *Educational Evaluation and Policy Analysis, 22*(2), 101-121.

Snijders, T. A. B. and R.J. Bosker. *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London: Sage Publications, 1999.

Somerville, M. (2011 February 4). When 'failing' schools close and whiter brighter schools open. *Huffington Post*. Retrieved from http://www.huffingtonpost.com/michele-somerville/when-failing-schools-clos_b_818181.html

Sorensen, J. & Goodman, L. (1997 June 13). Crew to shut 2 flunking schools. *New York Daily News.* Retrieved from http://articles.nydailynews.com/1997-06-13/news/18032299_1_chancellor-rudy-crew-schools-aides

Springer, M.G., Houck, E.A., Ceperley, P.E. and Hange, J. (2007). Revenue generation and resource allocation and deployment practices in smaller learning communities: Lessons learned from three high schools. *Journal of Education Finance, 32(4),* 443-469.

Stiefel, L., Bel Hadj Amor, H. & Schwartz, A. (2005). Best schools, worst schools, and school efficiency: A reconciliation and assessment of alternative classification systems. In Fowler, W.J., Jr. (Ed.). *Developments in school finance, 2004: Fiscal proceedings from the annual state data conference of July 2004* (NCES 2005-

865). U.S. Department of Education, National Center for Education Statistics. Washington, DC: Government Printing Office.

Stuart, E.A. (2007). Estimating causal effects using school-level datasets. *Educational Researcher, 36*(4), 187-198.

Stuart, E.A. (2010) Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1-12.

Stuit, D. (2010 December). *Are bad schools immortal?* Washington DC: Thomas B. Fordham Institute.

Summers, A., & Wolfe, B. (1977). Do schools make a difference? *American Economic Review, 67*, 639-52.

Supovitz, J. A. (2002). Developing communities of instructional practice. *Teachers College Record, 104*(8), 1591–1626.

Tucker, J. (2012 June 18). Families risk arrest in protesting school closure. San Francisco Chronicle. Retreived from http://www.sfgate.com/education/article/Families-risk-arrest-in-protesting-school-closure-3642840.php.

U.S. Department of Education (2010a Jan). *School Improvement Grants application.* Washington, D.C.: U.S. Department of Education.

U.S. Department of Education (2010b March). *ESEA blueprint for reform.* Washington, D.C.: U.S. Department of Education, Office of Planning, Evaluation and Policy Development.

U.S. Department of Education (2010c December). The state of school turnarounds. *School Turnaround* I(3). Retrieved from http://www2.ed.gov/programs/sif/sig1210newsltr.pdf

U.S. Department of Education (2011 March). *Smart ideas to increase educational productivity and student achievement.* Retrieved from: http://www.ed.gov/news/press-releases/department-education-provides-promising-practices-productivity-flexibility

U.S. Department of Education. (2008, April 3). *Peer review guidance for the NCLB differentiated accountability pilot applications*. Retrieved from http://www2.ed.gov/policy/elsec/guid/schoolimprovementguid.pdf

U.S. Department of Education. (2009). Overview Information; Race to the Top Fund; Notice Inviting Applications for New Awards for Fiscal Year (FY) 2010; Notice. *Federal Register* 74 (221), Pt 4.

United Federation of Teachers (2009). *Parents, teachers protest DOE plans to close Brooklyn school.* Retrieved from http://www.uft.org/news/issues/press/parents_teachers_protest/.

United States Government Accountability Office (April 2012). *School Improvement Grants: Education should take additional steps to enhance accountability for schools and contractors.* GAO 12-373. Washington, DC: Author.

U.S. Department of Education, Press Office. (2012). *Statement from U.S. Secretary of Education Arne Duncan on the new MDRC Study, 'Sustained Positive Effects on Graduation Rates Produced by New York City's Small Public High Schools of Choice'* [Press release]. Retrieved from http://www.ed.gov/news/press-releases/statement-us-secretary-education-arne-duncan-new-mdrc-study-sustained-positive-e

van der Laan, M. & Robins, J. (2003). *Unified methods for censored longitudinal data and causality.* New York: Springer.

Wallace, J.D. (2011 March 9) Carson students, parents, and teachers worry about closure. *KOLD News 13.* Retrieved from: http://www.kold.com/Global/story.asp?S=14220973

Walz, M., Arp, C. & Phillips, A. (2011 January 24). Scenes from three hearings. *GothamSchools.* Retrieved from: http://gothamschools.org/2011/01/24/scenes-from-three-hearings-jamaica-columbus-and-robeson/

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher, 37,* 469–479.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative science quarterly, 21*(1), 1-19.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of personnel evaluation in education, 11*, 57-68.

Young, V. M., Humphrey, D. C., Wang, H., Bosetti, K. R., Cassidy, L., Wechsler, M. E., . . . Schanzenbach, D. (2009). *Renaissance schools fund-supported schools: Early outcomes, challenges, and opportunities.* Chicago: Consortium on Chicago School Research.

Ziebarth, T. (2002). *State takeovers and reconstitutions. Policy Brief*. Denver, CO: Education Commission of the States.

Zimmer, R.W. and Toma, E.F. (2000). Peer effects in private and public schools across countries. *Journal of Policy Analysis and Management,* 19(1), 75-92.